GLOBAL REGIME FOR GENERAL ADDITIVE FUNCTIONALS OF CONDITIONED BIENAYMÉ-GALTON-WATSON TREES

Romain Abraham, Jean-François Delmas, Michel Nassif

GLOBAL REGIME FOR GENERAL ADDITIVE FUNCTIONALS OF CONDITIONED BIENAYMÉ-GALTON-WATSON TREES

ROMAIN ABRAHAM, JEAN-FRANÇOIS DELMAS, AND MICHEL NASSIF

Abstract. We give an invariance principle for very general additive functionals of conditioned Bienaymé-Galton-Watson trees in the global regime when the offspring distribution lies in the domain of attraction of a stable distribution, the limit being an additive functional of a stable Lévy tree. This includes the case when the offspring distribution has finite variance (the Lévy tree being then the Brownian tree). We also describe, using an integral test, a phase transition for toll functions depending on the size and height.

1. Introduction

In view of the many applications of trees (in computer science, biology, physics, ...), the study of additive functionals on large random trees has seen a lot of development in recent years, see references below. In this paper, we consider asymptotics for general additive functionals on conditioned Bienaymé-Galton-Watson (BGW for short) trees in the so-called global regime.

Recall that a functional F defined on finite rooted ordered discrete trees is said to be additive if it satisfies the recursion

$$F(\mathbf{t}) = \sum_{i=1}^{d} F(\mathbf{t}_i) + f(\mathbf{t}), \tag{1.1}$$

where $\mathbf{t}_1, \ldots, \mathbf{t}_d$ are the subtrees rooted at the d children of the root of the tree $\mathbf{t}$ and f is a given toll function. Notice that this can also be written as

$$F(\mathbf{t}) = \sum_{w\in\mathbf{t}} f(\mathbf{t}_w), \tag{1.2}$$

where $\mathbf{t}_w$ is the subtree of $\mathbf{t}$ above the vertex w and rooted at w. Such functionals are encountered in computer science where they represent the cost of divide-and-conquer algorithms, in phylogenetics where they are used as a rough measure of tree shape to detect imbalance or in chemical graph theory where they appear as a predictive tool for some chemical properties. Among these, we mention the total path length defined as the sum of the distances to the root of all vertices, the Wiener index [43] defined as the sum of the distances between all pairs of vertices, the shape functional, the Sackin index, the Colless index and the cophenetic index, see [42] for their definitions and also [14] for their representation using additive functionals, and the references therein. See also [39] for other functionals such that the number of matchings, dominating sets, independent

Date: September 16, 2020.
2010 *Mathematics Subject Classification.* 60J80, 60F17 05C05.
Key words and phrases. Galton-Watson trees, Lévy trees, additive functionals, scaling limit, phase transition.

sets for trees. We also mention the Shao and Sokal's B_1 index [6, 42] defined by

$$B_1(\mathbf{t}) = \sum_{\substack{w\in\mathbf{t}^\circ\\ w\neq\emptyset}} \frac{1}{\mathfrak{h}(\mathbf{t}_w)}, \tag{1.3}$$

where for every finite rooted ordered tree $\mathbf{t}$, $\mathfrak{h}(\mathbf{t})$ is its height and $\mathbf{t}^\circ$ is the set of internal vertices. It is used for assessing the balance of phylogenetic trees, see *e.g.* [21, 27, 31, 38, 41].

We shall consider in this paper random discrete trees τ^n which are BGW trees conditioned to have n vertices, and then study the limit of rescaled additive functionals as n goes to infinity. One can distinguish between local and global regime. In the local regime, the toll function is small or even vanishes when the subtree is large; so the main contribution to the additive functional comes from the small subtrees. These being almost independent, we understand intuitively why the limit distribution is Gaussian. See [29, 39, 45] for asymptotic results in the local regime. In the global regime, the toll function is large when the subtree is large; so the main contribution comes from large subtrees which are strongly dependent. This intuitively explain why we expect the limit to be non-Gaussian. As far as we know, asymptotic results in the global regime deal with toll functions depending only on the size. In this paper, we shall focus on the global regime for general toll functions. In particular, our results apply to toll functions depending on the size and height. When the toll function is monomial in the size of the tree $f(\mathbf{t}) = |\mathbf{t}|^{\alpha'}$, with $|\mathbf{t}|$ the cardinal of $\mathbf{t}$, Fill and Kapur [24] observed a phase transition at $\alpha' = 1/2$ for binary trees under the Catalan model (which is a special case of conditioned BGW trees): the global regime corresponds to $\alpha' > 1/2$. This was later generalized by Fill and Janson [23] to BGW trees with critical offspring distribution with finite variance using techniques from complex analysis; they identified a local regime for $\alpha' < 0$ and an intermediate regime for $0 < \alpha' < 1/2$. When the offspring distribution has infinite variance but lies in the domain of attraction of a stable distribution with index $\gamma \in (1, 2]$, Delmas, Dhersin and Sciauveau [14] proved convergence in distribution for $\alpha' \geq 1$ using stable Lévy trees and conjectured a phase transition at $\alpha' = 1/\gamma$. We shall prove this conjecture, as a particular case of our main result, see Theorem 1.1.

Let ξ be a $\mathbb{N}$-valued random variable. We write BGW(ξ) tree for a BGW tree with offspring distribution (the law of) ξ. We denote by τ^n a BGW(ξ) tree conditioned to have n vertices and we assume that ξ is critical, *i.e.* $\mathbb{E}[\xi] = 1$, nondegenerate, *i.e.* $\mathbb{P}(\xi = 0) > 0$, and that it belongs to the domain of attraction of a stable distribution with index $\gamma \in (1, 2]$, *i.e.* there exists a positive sequence $(b_n, n \geq 1)$ such that if $(\xi_n, n \geq 1)$ is a sequence of independent random variables with the same distribution as ξ then $b_n^{-1} \left(\sum_{k=1}^n \xi_k - n\right)$ converges in distribution towards a stable random variable whose Laplace transform is given by $\exp(\kappa\lambda^\gamma)$ for $\lambda \geq 0$, with index $\gamma \in (1, 2]$ and normalizing constant $\kappa > 0$ (the constant κ depends on the choice of the sequence $(b_n, n \geq 1)$. Under these assumptions, it is also well known that, as n goes to infinity, τ^n properly rescaled converges in distribution with respect to the Gromov-Hausdorff-Prokhorov topology to the stable Lévy tree $\mathcal{T}$ with index γ (and branching mechanism $\psi(\lambda) = \kappa\lambda^\gamma$) which is a rooted random real tree (see Section 4.2 for a precise definition), see Aldous [7] for the finite variance case and Duquesne [15] for the general case. The stable Lévy tree is a generalization of Aldous' Brownian continuum random tree which corresponds to $\gamma = 2$. We recall that the stable Lévy tree is the real tree coded by the normalized excursion of the height process associated with a stable Lévy process and that it codes the genealogy of continuous-state branching processes, see *e.g.* Le Gall and Le Jan [35], Duquesne and Le Gall [16, 17]. We recall that any real tree T is endowed with the length measure $\ell(\mathrm{d}y)$ (which roughly speaking is the Lebesgue measure on the branches of the tree) and that the Lévy tree is naturally endowed with a mass measure (which roughly speaking is the uniform probability measure on the infinite set of leaves). One of our main results can be stated as follows. We refer the reader to Proposition 7.1 and Theorem 7.3 for more general statements. Recall that $\mathbf{t}^\circ$ denotes the set of internal vertices of the discrete tree $\mathbf{t}$.

Theorem 1.1. *Let τ^n be a BGW(ξ) tree conditioned to have n vertices, with ξ being critical, nondegenerate and in the domain of attraction of a stable distribution with index $\gamma \in (1, 2]$. We suppose moreover that the sequence $(b_n,\, n \geq 1)$ defined as above is such that $(b_n/n^{1/\gamma},\, n \geq 1)$ is bounded away from zero and infinity. Let $\mathcal{T}$ be the stable Lévy tree with branching mechanism $\psi(\lambda) = \kappa\lambda^\gamma$. Let $\alpha', \beta \in \mathbb{R}$.*

(i) If $\gamma\alpha' + (\gamma - 1)\beta > 1$, we have the convergence in distribution and of the first moment

$$\frac{b_n^{1+\beta}}{n^{1+\alpha'+\beta}} \sum_{w \in \tau^{n,\circ}} |\tau_w^n|^{\alpha'}\, \mathfrak{h}(\tau_w^n)^\beta \xrightarrow[n\to\infty]{(d)+\text{mean}} \int_{\mathcal{T}} \mathfrak{m}(\mathcal{T}_y)^{\alpha'}\, \mathfrak{h}(\mathcal{T}_y)^\beta\, \ell(\mathrm{d}y), \tag{1.4}$$

where the right hand-side of (1.4) has finite mean and, for $y \in \mathcal{T}$, $\mathcal{T}_y$ is the subtree of $\mathcal{T}$ above y, $\mathfrak{m}(\mathcal{T}_y)$ is its mass, and $\mathfrak{h}(\mathcal{T}_y)$ its height.

(ii) If $\gamma\alpha' + (\gamma - 1)\beta \leq 1$, we have the convergence in distribution and of the first moment

$$\frac{b_n^{1+\beta}}{n^{1+\alpha'+\beta}} \sum_{w \in \tau^{n,\circ}} |\tau_w^n|^{\alpha'}\, \mathfrak{h}(\tau_w^n)^\beta \xrightarrow[n\to\infty]{(d)+\text{mean}} \infty. \tag{1.5}$$

We complete the previous result with some comments.

Remark 1.2.
(i) From Theorem 1.1, we obtain a phase change for functionals of the mass and height at $\gamma\alpha' + (\gamma-1)\beta = 1$. Heuristically, the condition on α' and β is due to the fact that the height of a (unnormalized) stable Lévy tree scales as its mass to the power $(\gamma-1)/\gamma$. Let us mention that this phase change is specific to BGW trees, see Remark 4.13 in this direction.
(ii) See conditions (ξ1) and (ξ2) in Section 4 for a more detailed discussion of the assumptions on the offspring distribution. The additional boundedness assumption on $(b_n/n^{1/\gamma},\, n \geq 1)$ is also equivalent to (ξ2)′. This latter can be dropped in (i) of Theorem 1.1 when $\alpha' \geq 1$ and $\beta \geq 0$ according to Proposition 4.10.
(iii) We also have the convergence (and finiteness) of the moments of all order $p > 1$ in (1.4) as soon as $p(\gamma\alpha + (\gamma-1)\beta) > 1 - \gamma$, with $\alpha = \alpha' - 1$, see Proposition 7.1. In particular for $\beta = 0$, we have the convergence of all nonnegative moments for $\alpha' \geq 1$. However, in the finite variance case, for $\alpha' \in (1/2, 1)$ (and $\beta = 0$), our result is not optimal, see (vi) below.
(iv) Theorem 1.1 generalizes a result by Delmas, Dhersin and Sciauveau where only functionals of the mass are considered (*i.e.* $\beta = 0$), see [14, Lemma 4.6]. In particular, we prove the conjecture stated therein: when $\beta = 0$, there is a phase transition at $\alpha' = 1/\gamma$ (the parameter α therein corresponds to $\alpha' - 1$ here). If we fix $\alpha' = 0$ and let β vary, the phase transition occurs at $\beta = 1/(\gamma - 1) \geq 1$. In particular, Shao and Sokal's B_1 index, which corresponds to $\alpha = 0$ and $\beta = -1$, lies in the local regime, whatever the value of the index γ and is therefore not covered by our results. See also (v) below.
(v) If the offspring distribution has finite variance $\sigma_\xi^2 \in (0, \infty)$, one can take $b_n = b\sqrt{n}$ in which case $\mathcal{T}$ is distributed as the Brownian continuum random tree with branching mechanism $\psi(\lambda) = \sigma_\xi^2\lambda^2/(2b^2)$. For $b = \sigma_\xi$, the contour process of $\mathcal{T}$ is a standard Brownian motion under its normalized excursion measure.

(vi) Assume that the offspring distribution has finite variance $\sigma_\xi^2 \in (0, \infty)$, which implies that $\gamma = 2$. We consider the asymptotics in the local regime of $\sum_{w\in\tau^{n,\circ}} |\tau_w^n|^{\alpha'}\, \mathfrak{h}(\tau_w^n)^\beta$, that is when $\alpha', \beta \in \mathbb{R}$ such that $2\alpha' + \beta < 0$. Denote by $F_{\alpha',\beta}$ the additive functional (1.2) associated with the toll function $f_{\alpha',\beta}(\mathbf{t}) = |\mathbf{t}|^{\alpha'} \mathfrak{h}(\mathbf{t})^\beta \mathbf{1}_{\{|\mathbf{t}|>1\}}$. By [29, Theorem 1.5] and Lemma 4.5, we have

$$\frac{F_{\alpha',\beta}(\tau^n) - n\mu}{\sqrt{n}} \xrightarrow[n\to\infty]{(d)} \mathcal{N}(0, \varsigma^2),$$

where μ, ς^2 are finite and given by $\mu = \mathbb{E}\left[f_{\alpha',\beta}(\tau)\right]$ and by $\varsigma^2 = 2\,\mathbb{E}\left[f_{\alpha',\beta}(\tau)\left(F_{\alpha',\beta}(\tau) - |\tau|\mu\right)\right] - \mathrm{Var}(f_{\alpha',\beta}(\tau)) - \mu^2/\sigma_\xi^2$, and τ is the corresponding unconditioned BGW tree. In particular, this covers Shao and Sokal's B_1 index (where $\alpha' = 0$ and $\beta = -1$). Notice that this leaves a gap for $0 \leq 2\alpha' + \beta \leq 1$. At least when $\beta = 0$, the situation is well understood. Fill and Janson [23] identify three different regimes: the global regime for $\alpha' > 1/2$, the local regime for $\alpha' < 0$ and an intermediate regime for $0 < \alpha' < 1/2$. The nontrivial asymptotic behavior of $F_{\alpha',\beta}(\tau^n)$ for $\gamma \in (1, 2)$ and $\gamma\alpha' + (\gamma - 1)\beta \leq 1$ (that is the non global regime in the non quadratic case) is an open question.

(vii) When τ^n is uniformly distributed among the set of full binary ordered trees with n vertices (which corresponds to a conditioned BGW(ξ) tree with $\mathbb{P}(\xi = 0) = \mathbb{P}(\xi = 2) = 1/2$), Fill and Kapur [24] studied the local and global regime when the toll function is a power of the size of the tree. Concerning the global regime, they showed the convergence in distribution, using the convergence of all positive moments in (1.4) for $\alpha' > 1/2$ and $\beta = 0$, see Eq. (3.14) and Proposition 3.5 therein. In that case, one can take $b_n = \sqrt{n}$ and $\mathcal{T}$ is the Brownian tree with branching mechanism $\psi(\lambda) = \lambda^2/2$. See also Fill and Janson [23] for general critical offspring distribution with finite variance. The explicit formula for the first moment of the right hand-side of (1.4) are given by the right hand-side of (1.12) with $\kappa = 1/2$ and $\alpha = \alpha' - 1$.

(viii) As an application, using (1.4), we obtain, when $\alpha' > 1/\gamma$, in Example 7.5 (with $\alpha' = \alpha + 1$) an asymptotic expansion in distribution for $b_n\, n^{-(1+\alpha')} \sum_{w\in\tau^{n,\circ}} |\tau_w^n|^{\alpha'} \log |\tau_w^n|$.

More generally, if one views a discrete tree as a real tree, then the left-hand side in (1.4) is related to the discrete length measure $\ell_n(\mathrm{d}y) = \sum_{w\in\tau^n} \delta_w(\mathrm{d}y)$ of τ^n (after rescaling by b_n/n). One way to interpret the result would be to say that the sequence of measures $\int_{\tau^n} \delta_{\tau_y^n}\, \ell_n(\mathrm{d}y)$ converges in distribution to $\int_{\mathcal{T}} \delta_{\mathcal{T}_y}\, \ell(\mathrm{d}y)$ in some sense. One might then hope to prove that the mapping $T \mapsto \int_T \delta_{T_y}\, \ell(\mathrm{d}y)$ is continuous on the space of compact real trees. This is not true however, see Remark 4.13, one problem being that the length measure is not finite in general. To overcome this difficulty, our approach, inspired by [14], consists in considering the length measure biased by the size of the subtree above y, thus penalizing small subtrees.

More precisely let $\mathbb{T}$ be the space of (equivalent classes of) weighted rooted compact real trees (*i.e.* the set of quadruplets $(T, \emptyset, d, \mu)$ where (T, d) is a real tree, $\emptyset$ is a distinguished vertex of T called the root, and the mass measure μ is a finite measure on T). We recall that the length measure ℓ on a real tree (T, d) has an intrinsic definition. For every $(T, \emptyset, d, \mu) \in \mathbb{T}$, we define a measure Ψ_T on $\mathbb{T} \times \mathbb{R}_+$ by: for every nonnegative measurable function f defined on $\mathbb{T} \times \mathbb{R}_+$,

$$\Psi_T(f) = \int_T \mu(T_y) f\left(T_y, H(y)\right)\, \ell(\mathrm{d}y), \tag{1.6}$$

where $H(y) = d(\emptyset, y)$ denotes the height of y (*i.e.* the distance to the root) in T. We also consider the measure $\Psi_T^{\mathfrak{m}\mathfrak{h}}$ on $\mathbb{R}_+^2$ defined similarly to Ψ_T for functions depending only on the mass and height of the tree, see (3.2).

If $\mathbf{t}$ is a finite rooted ordered tree and $a > 0$, we denote by $a\mathbf{t}$ the real tree associated with $\mathbf{t}$, rescaled so that all edges have length a and equipped with the uniform probability measure on the set of vertices whose height is an integer multiple of a, see Section 2.3 for a precise definition. Furthermore, for $w \in \mathbf{t}$, we write aw for the corresponding vertex in $a\mathbf{t}$ and $a\mathbf{t}_w$ for the subtree of $a\mathbf{t}$ above aw. The height of w in $\mathbf{t}$ is denoted by $H(w)$; and thus the height of aw in $a\mathbf{t}$ is $aH(w)$. In the spirit of [14], we consider the measure $\mathcal{A}^\circ_{\mathbf{t},a}$ on $\mathbb{T} \times \mathbb{R}_+$ defined by: for nonnegative measurable function f defined on $\mathbb{T} \times \mathbb{R}_+$,

$$\mathcal{A}^\circ_{\mathbf{t},a}(f) = \frac{a}{|\mathbf{t}|} \sum_{w \in \mathbf{t}^\circ} |\mathbf{t}_w| f\left(a\mathbf{t}_w, aH(w)\right). \tag{1.7}$$

In (1.7), instead of summing over all the internal vertices ($w \in \mathbf{t}^\circ$) one could also sum over all vertices including the leaves ($w \in \mathbf{t}$); in this case the measure is denoted by $\mathcal{A}_{\mathbf{t},a}$. The two measures are close in total variation as $d_{\mathrm{TV}}(\mathcal{A}_{\mathbf{t},a}, \mathcal{A}^\circ_{\mathbf{t},a}) \leq a$, see (4.18). We mention that the measure $\mathcal{A}_{\mathbf{t},a}$ was already considered in [14] for functions f depending only on the size.

For every finite rooted ordered tree $\mathbf{t}$ and $a > 0$, we show (see Lemma 4.8) that the measures $\mathcal{A}^\circ_{\mathbf{t},a}$ and $\mathcal{A}_{\mathbf{t},a}$ can be approximated by $\Psi_{a\mathbf{t}}$. In Proposition 3.4, we give another expression for Ψ_T:

$$\Psi_T(f) = \int_T \mu(\mathrm{d}x) \int_0^{H(x)} f\left(T_{r,x}, r\right) \mathrm{d}r, \tag{1.8}$$

for every nonnegative measurable function f defined on $\mathbb{T} \times \mathbb{R}_+$. Here $T_{r,x}$ is the subtree of T above level r containing x. This latter expression of Ψ_T is used to prove it is continuous as a function of T, see Proposition 3.3.

Theorem 1.3. *The mapping $T \mapsto \Psi_T$, from $\mathbb{T}$ endowed with the Gromov-Hausdorff-Prokhorov topology to $\mathcal{M}(\mathbb{T} \times \mathbb{R}_+)$, the space of nonnegative finite measures on $\mathbb{T} \times \mathbb{R}_+$, endowed with the topology of weak convergence, is well defined and continuous.*

This allows to derive a general invariance principle: for any sequence of random discrete trees $(\tau^n,\, n \in \mathbb{N})$ such that $a_n \tau^n$ converges in distribution to some random real tree $\mathcal{T}$ in the Gromov-Hausdorff-Prokhorov topology where $(a_n,\, n \in \mathbb{N})$ is a sequence of positive numbers converging to 0 and such that $\left(a_n \mathbb{E}\left[\mathfrak{h}(\tau^n)\right],\, n \in \mathbb{N}\right)$ is bounded, one has the convergence in distribution of the measures $\mathcal{A}^\circ_{\tau^n,a_n}$ and $\mathcal{A}_{\tau^n,a_n}$ to $\Psi_{\mathcal{T}}$ (this is a consequence of Lemma 4.8 and Theorem 1.3). For example, this applies to Pólya trees, see Remark 4.12, which were shown to converge to the Brownian tree, see [25] and [36]. For BGW trees, we have the following result which is a direct consequence of the convergence on conditioned BGW trees to stable Lévy tree, see [15], and Theorem 1.3 and Lemma 4.8.

Corollary 1.4. *Let τ^n be a BGW(ξ) tree conditioned to have n vertices, with ξ satisfying (ξ1) and (ξ2), and $(b_n,\, n \geq 1)$ be defined as in Theorem 1.1. Let $\mathcal{T}$ be the stable Lévy tree with branching mechanism $\psi(\lambda) = \kappa\lambda^\gamma$. We have the following convergence in distribution and of all positive moments*

$$\frac{b_n}{n^2} \sum_{w \in \tau^{n,\circ}} |\tau^n_w| f\left(\frac{b_n}{n}\tau^n_w, \frac{b_n}{n} H(w)\right) \xrightarrow[n\to\infty]{(d)+\text{moments}} \Psi_{\mathcal{T}}(f),$$

where f is a bounded continuous real-valued function defined on $\mathbb{T} \times \mathbb{R}_+$.

We improve this result by allowing the function f to blow up as either the mass or the height goes to zero under the stronger assumption $(\xi 2)'$: see Proposition 7.1, and more precisely Theorem 7.3 when f is a product of a function of the mass and a function of the height, one of them being a power function. As a particular case, property (i) of Theorem 1.1 gives a precise result when f is a power function of the mass and the height. Related to this latter result, we give a complete description of the finiteness of $\Psi^{\mathfrak{mh}}_{\mathcal{T}}(f)$ for power functions f where $\mathcal{T}$ is the stable Lévy tree and we also compute its first moment. We refer to Corollaries 6.4 and 6.7, and Proposition 6.9 for a more general statement. By convention, we write $\Psi^{\mathfrak{mh}}_{\mathcal{T}}(g(x)h(u))$ for $\Psi^{\mathfrak{mh}}_{\mathcal{T}}(f)$ where $f(x,u) = g(x)h(u)$ and we see g as a function of the mass and h as a function of the height. In particular, thanks to (1.6), we have for $\alpha, \beta \in \mathbb{R}$ that $\Psi^{\mathfrak{mh}}_{\mathcal{T}}(x^\alpha u^\beta) = \int_{\mathcal{T}} \mathfrak{m}(\mathcal{T}_y)^{\alpha'}\, \mathfrak{h}(\mathcal{T}_y)^\beta\, \ell(\mathrm{d}y)$ with $\alpha' = \alpha + 1$.

Proposition 1.5. *Let $\mathcal{T}$ be the stable Lévy tree with branching mechanism $\psi(\lambda) = \kappa\lambda^\gamma$ and let $\alpha, \beta \in \mathbb{R}$. We have*

$$\gamma\alpha + (\gamma-1)(\beta+1) > 0 \iff \Psi^{\mathfrak{mh}}_{\mathcal{T}}(x^\alpha u^\beta) < \infty \ \text{a.s.} \iff \mathbb{E}\left[\Psi^{\mathfrak{mh}}_{\mathcal{T}}(x^\alpha u^\beta)\right] < \infty, \tag{1.9}$$

$$\gamma\alpha + (\gamma-1)(\beta+1) \leq 0 \iff \Psi^{\mathfrak{mh}}_{\mathcal{T}}(x^\alpha u^\beta) = \infty \ \text{a.s.} \iff \mathbb{E}\left[\Psi^{\mathfrak{mh}}_{\mathcal{T}}(x^\alpha u^\beta)\right] = \infty. \tag{1.10}$$

For every $\alpha, \beta \in \mathbb{R}$ such that $\gamma\alpha + (\gamma-1)(\beta+1) > 0$, we have

$$\mathbb{E}\left[\Psi^{\mathfrak{mh}}_{\mathcal{T}}(x^\alpha u^\beta)\right] = \frac{1}{\kappa^{1/\gamma}|\Gamma(-1/\gamma)|}\mathrm{B}(\alpha + (\beta+1)(1-1/\gamma), 1-1/\gamma)\, \mathbb{E}\left[\mathfrak{h}(\mathcal{T})^\beta\right], \tag{1.11}$$

where Γ is the gamma function and B is the beta function. Furthermore, we have $\mathbb{E}\left[\Psi^{\mathfrak{mh}}_{\mathcal{T}}(x^\alpha u^\beta)^p\right] < \infty$ for every $p \geq 1$ such that $p(\gamma\alpha + (\gamma-1)\beta) > 1-\gamma$. In the Brownian case ($\gamma = 2$), for every $\alpha, \beta \in \mathbb{R}$ such that $2\alpha + \beta + 1 > 0$, we have

$$\mathbb{E}\left[\Psi^{\mathfrak{mh}}_{\mathcal{T}}(x^\alpha u^\beta)\right] = \frac{1}{\sqrt{\pi\kappa}}\left(\frac{\pi}{\kappa}\right)^{\beta/2}\xi(\beta)\mathrm{B}\left(\alpha + \frac{\beta+1}{2}, \frac{1}{2}\right), \tag{1.12}$$

where ξ is the Riemann xi function defined by $\xi(s) = \frac{1}{2}s(s-1)\pi^{-s/2}\Gamma(s/2)\zeta(s)$ for every $s \in \mathbb{C}$ and ζ is the Riemann zeta function.

Thanks to Duquesne and Wang [18], $\mathbb{E}\left[\mathfrak{h}(\mathcal{T})^\beta\right]$ is finite for all $\beta \in \mathbb{R}$, so that the right hand side of (1.11) is finite.

We conclude the introduction by giving a formula for the distribution of $\mathcal{T}_y$, the subtree above y, when y is chosen according to the length measure $\ell(\mathrm{d}y)$ on the stable Lévy tree $\mathcal{T}$, see Proposition 6.3. This is a key result for the proof of Proposition 1.5 and it is also interesting by itself (it is in particular related to the additive coalescent and the uniform pruning on the skeleton of the Lévy tree, see Remark 6.2 in this direction). Let $\mathbf{N}$ denote the excursion measure of height process H which codes the (unnormalized) stable Lévy tree $\mathcal{T}_H$. (Notice that $\mathcal{T}$ under $\mathbb{P}$ is distributed as $\mathcal{T}_H$ conditionally on $\{\mathfrak{m}(\mathcal{T}_H) = 1\}$ under $\mathbf{N}$.)

Proposition 1.6. *Let $\mathcal{T}$ be the stable Lévy tree with branching mechanism $\psi(\lambda) = \kappa\lambda^\gamma$ where $\kappa > 0$ and $\gamma \in (1, 2]$. Let f be a nonnegative measurable function defined on $\mathbb{T}$. We have:*

$$\mathbb{E}\left[\int_{\mathcal{T}} f(\mathcal{T}_y)\, \ell(\mathrm{d}y)\right] = \mathbf{N}\left[(1 - \mathfrak{m}(\mathcal{T}_H))^{-1/\gamma}\mathbf{1}_{\{\mathfrak{m}(\mathcal{T}_H)<1\}}\, f(\mathcal{T}_H)\right].$$

The paper is organized as follows. Section 2 establishes notation and defines the main objects used in this paper (discrete trees using Neveu's formalism, real trees, Gromov-Hausdorff-Prokhorov topology). In Section 3, we give properties of the measure Ψ_T and prove its continuity with respect to T. Section 4 introduces the setting of BGW trees and stable Lévy trees and gives a first convergence result for continuous functions. We gather some technical results in Section 5. Section 6 is devoted to the study of functionals of the mass and height on the stable Lévy tree and Section 7 presents the general convergence result for functions that may blow up and describes the phase change. Appendix A introduces a space of measures and studies random elements thereof; its results are used in the proofs of Proposition 7.1 and Theorem 7.3.

2. Definitions and notations

2.1. **Weak convergence in a Polish space.** Let (S,ρ) be a Polish metric space. We denote by $\mathcal{B}(S)$ (resp. $\mathcal{B}_+(S)$, resp. $\mathcal{B}_b(S)$) the set of measurable functions defined on S and taking values in $[-\infty,+\infty]$ (resp. in $[0,+\infty]$, resp. in $\mathbb{R}$ and bounded) and by $\mathcal{C}(S)$ (resp. $\mathcal{C}_+(S)$, resp. $\mathcal{C}_b(S)$) the set of continuous real-valued functions defined on S (resp. nonnegative, resp. bounded). For $f\in\mathcal{B}(S)$, we set $\|f\|_\infty=\sup_{x\in S}|f(x)|$. For $f\in\mathcal{C}_b(S)$, we define its Lipschitz and bounded Lipschitz norm:

$$\|f\|_{\mathrm{L}}=\sup_{x\neq y}\frac{|f(x)-f(y)|}{\rho(x,y)}\quad\text{and}\quad \|f\|_{\mathrm{BL}}=\|f\|_\infty+\|f\|_{\mathrm{L}}\,.$$

We denote by $\mathcal{M}(S)$ the set of nonnegative finite measures on S. For every $\mu\in\mathcal{M}(S)$ and $f\in\mathcal{B}_+(S)$, we write $\mu(f)=\int f(x)\,\mu(\mathrm{d}x)$. The set $\mathcal{M}(S)$ is endowed with the topology of weak convergence which can be metrized (see [12, Section 8.3 and Theorem 8.3.2]) by the bounded Lipschitz distance (also known as the Kantorovich-Rubinstein distance): if $\mu,\nu\in\mathcal{M}(S)$, set

$$d_{\mathrm{BL}}(\mu,\nu)=\sup\left\{|\mu(f)-\nu(f)|,\ f\in\mathcal{C}_b(S)\text{ such that }\|f\|_{\mathrm{BL}}\leq 1\right\}.$$

Moreover, the space $(\mathcal{M}(S),d_{\mathrm{BL}})$ is Polish by [12, Theorem 8.9.4]. We also recall the total variation norm given by

$$d_{\mathrm{TV}}(\mu,\nu)=\frac{1}{2}\sup\left\{|\mu(f)-\nu(f)|,\ f\in\mathcal{B}(S)\text{ such that }\|f\|_\infty\leq 1\right\}.$$

2.2. **Discrete trees.** We recall Neveu's formalism for rooted ordered discrete trees. Let $\mathcal{U}=\cup_{n\geq 0}(\mathbb{N}^*)^n$ be the set of labels with the convention $(\mathbb{N}^*)^0=\{\emptyset\}$. If $v=(v^1,\ldots,v^n)\in\mathcal{U}$, we denote by $H(v)=n$. By convention, we set $H(\emptyset)=0$. If $v=(v^1,\ldots,v^n),w=(w^1,\ldots,w^m)\in\mathcal{U}$, we write $vw=(v^1,\ldots,v^n,w^1,\ldots,w^m)$ for the concatenation of v and w. In particular, $v\emptyset=\emptyset v=v$. We say that v is an ancestor of w and write $v\preccurlyeq w$ if there exists $u\in\mathcal{U}$ such that $w=vu$. If $v\preccurlyeq w$ and $v\neq w$ then we shall write $v\prec w$. The mapping $\mathrm{pr}\colon\mathcal{U}\setminus\{\emptyset\}\to\mathcal{U}$ is defined by $\mathrm{pr}(v^1,\ldots,v^n)=(v^1,\ldots,v^{n-1})$ ($\mathrm{pr}(v)$ is the parent of v). A finite rooted ordered tree $\mathbf{t}$ is a finite subset of $\mathcal{U}$ such that

(i) $\emptyset\in\mathbf{t}$,
(ii) $v\in\mathbf{t}\setminus\{\emptyset\}\Rightarrow\mathrm{pr}(v)\in\mathbf{t}$,
(iii) for every $v\in\mathbf{t}$, there exists a finite integer $k_v(\mathbf{t})\geq 0$ such that, for every $j\in\mathbb{N}^*$, $vj\in\mathbf{t}$ if and only if $1\leq j\leq k_u(\mathbf{t})$.

The number $k_v(\mathbf{t})$ is interpreted as the number of children of the vertex v in $\mathbf{t}$, $H(v)$ is its generation, $\mathrm{pr}(v)$ is its parent and more generally, the vertices $v,\mathrm{pr}(v),\mathrm{pr}^2(v),\ldots\mathrm{pr}^{H(v)}(v)=\emptyset$ are its ancestors.

The vertex v is called a leaf (resp. internal vertex) if $k_v(\mathbf{t}) = 0$ (resp. $k_v(\mathbf{t}) > 0$). The vertex $\emptyset$ is called the root of $\mathbf{t}$. We denote the set of leaves by $\mathrm{Lf}(\mathbf{t})$ and the set of internal vertices by $\mathbf{t}^\circ$. If $v \in \mathbf{t}$, we define the subtree $\mathbf{t}_v$ of $\mathbf{t}$ above v as

$$\mathbf{t}_v = \{w \in \mathcal{U} \colon \ vw \in \mathbf{t}\}.$$

Moreover, for every $0 \leq k \leq H(v)$, we define the subtree $\mathbf{t}_{k,\,v}$ of $\mathbf{t}$ above level k containing v as

$$\mathbf{t}_{k,\,v} = \mathbf{t}_{\mathrm{pr}^{H(v)-k}(v)}$$

where $\mathrm{pr}^{H(v)-k}(v)$ is the unique ancestor of v with height k, with the convention that $\mathrm{pr}^0(v) = v$. We denote by $|\mathbf{t}| = \mathrm{Card}(\mathbf{t})$ the number of vertices of $\mathbf{t}$ and by $\mathfrak{h}(\mathbf{t}) = \sup_{v\in\mathbf{t}} H(v)$ the height of $\mathbf{t}$.

2.3. **Real trees.** We recall the formalism of real trees, see [20]. A metric space (T, d) is a real tree if the following two properties hold for every $x, y \in T$.

(i) (Unique geodesics). There exists a unique isometric map $f_{x,y} \colon [0, d(x, y)] \to T$ such that $f_{x,y}(0) = x$ and $f_{x,y}(d(x, y)) = y$.
(ii) (Loop-free). If φ is a continuous injective map from $[0, 1]$ into T such that $\varphi(0) = x$ and $\varphi(1) = y$, then we have $\varphi([0, 1]) = f_{x,y}\left([0, d(x, y)]\right)$.

For a rooted real tree $(T, \emptyset, d)$, that is a real tree with a distinguished vertex $\emptyset \in T$ called the root, we define the set of leaves by

$$\mathrm{Lf}(T) = \{x \in T \setminus \{\emptyset\} \colon \ T \setminus \{x\} \text{ is connected}\}\,,$$

with the convention that $\mathrm{Lf}(T) = \{\emptyset\}$ if $T = \{\emptyset\}$. A weighted rooted real tree $(T, \emptyset, d, \mu)$ is a rooted real tree $(T, \emptyset, d)$ equipped with a nonnegative finite measure μ. In what follows, real trees will always be weighted and rooted and we will simply call them real trees.

Let us consider a real tree $(T, \emptyset, d, \mu)$. The total mass of the tree T is defined by $\mathfrak{m}(T) = \mu(T)$ and its height by $\mathfrak{h}(T) = \sup_{x\in T} H(x) \in [0, \infty]$, with $H(x) = d(\emptyset, x)$ the height of x. Note that if (T, d) is compact, then $\mathfrak{h}(T) < \infty$. The range of the mapping $f_{x,y}$ described in (i) above is denoted by $[\![x, y]\!]$ (this is the line segment between x and y in the tree). We also write $[\![x, y[\![= [\![x, y]\!] \setminus \{y\}$. In particular, $[\![\emptyset, x]\!]$ is the path going from the root to x which we will interpret as the ancestral line of vertex x. We define a partial order on the tree by setting $x \preccurlyeq y$ (x is an ancestor of y) if and only if $x \in [\![\emptyset, y]\!]$. If $x, y \in T$, there is a unique $z \in T$ such that $[\![\emptyset, x]\!] \cap [\![\emptyset, y]\!] = [\![\emptyset, z]\!]$. We write $z = x \wedge y$ and call it the most recent common ancestor of x and y. Let $x \in T$ be a vertex. Let $r \in [0, H(x)]$. We denote by $x_r \in T$ be the unique ancestor of x with height $H(x_r) = r$. As in the discrete case, we also define the subtree T_x of T above x as

$$T_x = \{y \in T \colon \ x \preccurlyeq y\}\,,$$

and the subtree $T_{r,\,x} = T_{x_r}$ of T above level r containing x as

$$T_{r,\,x} = \{y \in T \colon \ H(x \wedge y) \geq r\} = T_{x_r}.$$

Then T_x (resp. $T_{r,\,x}$) can be naturally viewed as a real tree, rooted at x (resp. at x_r) and endowed with the distance d and the measure $\mu_{|T_x} = \mu(\cdot \cap T_x)$ (resp. the measure $\mu_{|T_{r,\,x}}$). Note that $T_{0,\,x} = T$ and $T_{H(x),\,x} = T_x$.

Remark 2.1. We recall the construction of a real tree from an excursion path, see *e.g.* [20, Example 3.14] or [17, Section 2.1]. Let e be a positive excursion path, that is $e \in \mathcal{C}_+(\mathbb{R}_+)$ such that $e(0) = 0$, $e(s) > 0$ for $0 < s < \sigma$ and $e(s) = 0$ for $s \geq \sigma$ where $\sigma \coloneqq \inf\{s > 0\colon\ e(s) = 0\} \in (0, \infty)$ is the duration of the excursion. Set $d_e(t, s) = e(t) + e(s) - 2 \inf_{[t\wedge s, t\vee s]} e$ for every $t, s \in [0, \sigma]$ and define an equivalence relation on $[0, \sigma]$ by letting $t \sim_e s$ if and only if $d_e(t, s) = 0$. The real tree T_e coded by e is defined as the quotient space $[0, \sigma]/\sim_e$ rooted at $p(0)$ where $p\colon [0, \sigma] \to T_e$ is the quotient map and equipped with the distance d_e and the pushforward measure $\lambda \circ p^{-1}$ where λ is the Lebesgue measure on $[0, \sigma]$. This defines a compact weighted rooted real tree. Notice that the mass and height of T_e are given by $\mathfrak{m}(T_e) = \sigma$ and $\mathfrak{h}(T_e) = \|e\|_\infty$.

We will need to view discrete trees as real trees. Let $\mathbf{t}$ be a finite rooted ordered tree and let $a > 0$. Suppose that $\mathbf{t}$ is embedded into the plane such that the edges are straight lines with length a that only intersect at their incident vertices. Denote by $\pi_{\mathbf{t},a}\colon \mathbf{t} \to \mathbb{R}^2$ the embedding and by $a\mathbf{t} = \pi_{\mathbf{t},a}(\mathbf{t}) \subset \mathbb{R}^2$ the embedded set. Moreover, for a vertex $v \in \mathbf{t}$, we denote by $av = \pi_{\mathbf{t},a}(v)$ the corresponding vertex in $a\mathbf{t}$. Then $a\mathbf{t}$ can be considered as a compact real tree $(a\mathbf{t}, d_{\mathbf{t}}, \mu_{\mathbf{t}})$: the distance $d_{\mathbf{t}}(x, y)$ between two points $x, y \in a\mathbf{t}$ is defined as the shortest length of a curve that connects x and y, and the measure $\mu_{\mathbf{t}}$ is the pushforward of the uniform probability measure on $\mathbf{t}$ by the embedding $\pi_{\mathbf{t},a}$. In other words, $a\mathbf{t}$ is obtained from $\mathbf{t}$ by connecting every vertex to its children in such a way that all edges have length a and is equipped with the measure $\mu_{\mathbf{t}}$ supported on the set $\{av\colon\ v \in \mathbf{t}\}$ and satisfying $\mu_{\mathbf{t}}(\{av\}) = 1/|\mathbf{t}|$ for every $v \in \mathbf{t}$. The tree $a\mathbf{t}$ is naturally rooted at $a\emptyset$ (also denoted $\emptyset$). Notice that vertices of the form av with $v \in \mathbf{t}$ are precisely those vertices in $a\mathbf{t}$ whose height is an integer multiple of a. Finally, to simplify notation, for every $v \in \mathbf{t}$, we will write $a\mathbf{t}_v$ instead of $(a\mathbf{t})_{av}$ for the subtree of $a\mathbf{t}$ above av. We stress that, unless $v = \emptyset$, the measure of the compact real tree $a\mathbf{t}_v$ has mass less than one, whereas the measure of the compact real tree $a(\mathbf{t}_v)$ is by definition a probability measure.

2.4. **Gromov-Hausdorff-Prokhorov topology.** Denote by $\mathbb{T}$ the set of measure-preserving and root-preserving isometry classes of compact real trees. We will often identify a class with an element of this class. So we shall write that $(T, \emptyset, d, \mu) \in \mathbb{T}$ if $(T, \emptyset, d)$ is a rooted compact real tree and μ is a nonnegative finite measure on T. When there is no ambiguity, we may write T for $(T, \emptyset, d, \mu)$.

We start by giving the standard definition of the Gromov-Hausdorff-Prokhorov distance. Let (E, δ) be a metric space. Given a non-empty subset $A \subset E$ and $\varepsilon > 0$, the ε-neighborhood of A is $A^\varepsilon = \{x \in E\colon\ d(x, A) < \varepsilon\}$. The Hausdorff distance δ_{H} between two non-empty subsets $A, B \subset E$ is defined by

$$\delta_{\mathrm{H}}(A, B) = \inf\{\varepsilon > 0\colon\ A \subset B^\varepsilon \text{ and } B \subset A^\varepsilon\}.$$

Next, denoting by $\mathcal{B}(E)$ the Borel σ-field on (E, δ), the Lévy-Prokhorov distance between two finite nonnegative measures μ, ν on $(E, \mathcal{B}(E))$ is

$$\delta_{\mathrm{P}}(\mu, \nu) = \inf\{\varepsilon > 0\colon\ \mu(A) \leq \nu(A^\varepsilon) + \varepsilon \text{ and } \nu(A) \leq \mu(A^\varepsilon) + \varepsilon,\ \forall A \in \mathcal{B}(E)\}.$$

We can now give the standard distance used to define the Gromov-Hausdorff-Prokhorov topology. For two compact real trees $(T, \emptyset, d, \mu), (T', \emptyset', d', \mu') \in \mathbb{T}$, set

$$d^\circ_{\mathrm{GHP}}(T, T') = \inf\left\{\delta(\varphi(\emptyset), \varphi'(\emptyset')) \vee \delta_{\mathrm{H}}(\varphi(T), \varphi'(T')) \vee \delta_{\mathrm{P}}(\mu \circ \varphi^{-1}, \mu' \circ \varphi'^{-1})\right\}, \tag{2.1}$$

where the infimum is taken over all isometries $\varphi\colon T \to E$ and $\varphi'\colon T' \to E$ into a common metric space (E, δ). This defines a metric which induces the Gromov-Hausdorff-Prokhorov topology on $\mathbb{T}$.

It will be convenient for our purposes to define another metric which induces the same topology on $\mathbb{T}$. Let $(T,\emptyset,d,\mu),(T',\emptyset',d',\mu')\in\mathbb{T}$. Recall that a correspondence between T and T' is a subset $\mathcal{R}\subset T\times T'$ such that for every $x\in T$, there exists $x'\in T'$ such that $(x,x')\in\mathcal{R}$, and conversely, for every $x'\in T'$, there exists $x\in T$ such that $(x,x')\in\mathcal{R}$. In other words, if we denote by $p\colon T\times T'\to T$ (resp. $p'\colon T\times T'\to T'$) the canonical projection on T (resp. on T'), a correspondence is a subset $\mathcal{R}\subset T\times T'$ such that $p(\mathcal{R})=T$ and $p'(\mathcal{R})=T'$. If $\mathcal{R}$ is a correspondence between T and T', its distortion is defined by

$$\mathrm{dis}(\mathcal{R})=\sup\left\{|d(x,y)-d'(x',y')|:\ (x,x'),(y,y')\in\mathcal{R}\right\}.$$

Next, for any nonnegative finite measure m on $T\times T'$, we define its discrepancy with respect to μ and μ' by

$$\mathrm{D}(m;\mu,\mu')=d_{\mathrm{TV}}(m\circ p^{-1},\mu)+d_{\mathrm{TV}}(m\circ p'^{-1},\mu').$$

Then the Gromov-Hausdorff-Prokhorov distance between T and T' is defined as

$$d_{\mathrm{GHP}}(T,T')=\inf\left\{\frac{1}{2}\,\mathrm{dis}(\mathcal{R})\vee\mathrm{D}(m;\mu,\mu')\vee m(\mathcal{R}^c)\right\},\tag{2.2}$$

where the infimum is taken over all correspondences $\mathcal{R}$ between T and T' such that $(\emptyset,\emptyset')\in\mathcal{R}$ and all nonnegative finite measures m on $T\times T'$. It can be verified that d_{GHP} is indeed a distance on $\mathbb{T}$ which is equivalent to d°_{GHP} and that the space $(\mathbb{T},d_{\mathrm{GHP}})$ is a Polish metric space, see [4].

We gather some facts about the Gromov-Hausdorff-Prokhorov distance that will be useful later. We refer the reader to [4] or [40]. We have that

$$\frac{1}{2}\,|\mathfrak{h}(T)-\mathfrak{h}(T')|\vee|\mathfrak{m}(T)-\mathfrak{m}(T')|\leq d_{\mathrm{GHP}}(T,T')\leq(\mathfrak{h}(T)+\mathfrak{h}(T'))\vee(\mathfrak{m}(T)+\mathfrak{m}(T')).\tag{2.3}$$

When $T'=\{\emptyset\}$ is the trivial tree consisting only of the root with mass 0, we have

$$\frac{1}{2}\mathfrak{h}(T)\vee\mathfrak{m}(T)\leq d_{\mathrm{GHP}}(T,\{\emptyset\})\leq\mathfrak{h}(T)\vee\mathfrak{m}(T).\tag{2.4}$$

We consider the subset of $\mathbb{T}$ of trees with either height or mass equal to 0:

$$\mathbb{T}_0=\{T\in\mathbb{T}\colon\ \mathfrak{m}(T)=0\text{ or }\mathfrak{h}(T)=0\}.\tag{2.5}$$

Note that $\mathbb{T}_0\subset\mathbb{T}$ is a closed subset since the mappings $\mathfrak{m}\colon\mathbb{T}\to\mathbb{R}$ and $\mathfrak{h}\colon\mathbb{T}\to\mathbb{R}$ are continuous with respect to the Gromov-Hausdorff-Prokhorov topology, thanks to (2.3). We now give bounds for the distance of a tree T to $\mathbb{T}_0$ which are similar to (2.4).

Lemma 2.2. *Let $T\in\mathbb{T}$. Then we have*

$$\frac{1}{2}\mathfrak{h}(T)\wedge\mathfrak{m}(T)\leq d_{\mathrm{GHP}}(T,\mathbb{T}_0)\leq\mathfrak{h}(T)\wedge\mathfrak{m}(T).\tag{2.6}$$

Proof. Let $(T,d,\emptyset,\mu)\in\mathbb{T}$ and $\delta>d_{\mathrm{GHP}}(T,\mathbb{T}_0)$. Then there exists $T'\in\mathbb{T}_0$ such that $d_{\mathrm{GHP}}(T,T')\leq\delta$. By (2.3), we get

$$\frac{1}{2}\,|\mathfrak{h}(T)-\mathfrak{h}(T')|\vee|\mathfrak{m}(T)-\mathfrak{m}(T')|\leq\delta.$$

But since $T'\in\mathbb{T}_0$, either $\mathfrak{h}(T')=0$ or $\mathfrak{m}(T')=0$. Therefore, either $\mathfrak{h}(T)\leq 2\delta$ or $\mathfrak{m}(T)\leq\delta$. Since $\delta>d_{\mathrm{GHP}}(T,\mathbb{T}_0)$ is arbitrary, this yields the lower bound.

To prove the upper bound, let $T' = T$ endowed with the zero measure $\mu' = 0$, and take $\mathcal{R} = \{(x,x)\colon x \in T\}$ and m the zero measure on $T \times T'$. Then $\mathrm{dis}(\mathcal{R}) = 0$, $m(\mathcal{R}^c) = 0$ and $\mathrm{D}(m;\mu,\mu') = \mu(T) = \mathfrak{m}(T)$. It follows that $d_{\mathrm{GHP}}(T,T') \leq \mathfrak{m}(T)$. Note that $T' \in \mathbb{T}_0$, therefore

$$d_{\mathrm{GHP}}(T,\mathbb{T}_0) \leq d_{\mathrm{GHP}}(T,T') \leq \mathfrak{m}(T).$$

Next, let $T'' = \{\emptyset\}$ be the trivial tree consisting only of the root with mass $\mathfrak{m}(T)$, *i.e.* endowed with the measure $\mu'' = \mathfrak{m}(T)\delta_\emptyset$. Take $\mathcal{R} = T \times \{\emptyset\}$ and $m(A \times B) = \mu(A)\delta_\emptyset(B)$. Then, we have $\mathcal{R}^c = \emptyset$, so $m(\mathcal{R}^c) = 0$. Moreover, we have

$$\mathrm{dis}(\mathcal{R}) = \sup\{|d(x,y)|\colon x,y \in T\} \leq 2\mathfrak{h}(T).$$

Since $m \circ p^{-1} = \mu$ and $m \circ p''^{-1} = \mathfrak{m}(T)\delta_\emptyset = \mu''$, we get $\mathrm{D}(m,\mu,\mu'') = 0$. It follows that $d_{\mathrm{GHP}}(T,T'') \leq \mathfrak{h}(T)$. Since $T'' \in \mathbb{T}_0$, we deduce that

$$d_{\mathrm{GHP}}(T,\mathbb{T}_0) \leq d_{\mathrm{GHP}}(T,T'') \leq \mathfrak{h}(T).$$

This finishes the proof of the upper bound. □

3. A finite measure indexed by a tree

Let $(T,\emptyset,d,\mu)$ be a compact real tree. Let $x \in T$ and $r \in [0,H(x)]$, where $H(x) = d(\emptyset,x)$. Recall that $T_{r,x} = \{y \in T : H(x \wedge y) \geq r\}$ is the subtree containing x and starting at height r, endowed with the distance d and the measure $\mu_{|T_{r,x}}$. It is straightforward to check that $T_{r,x}$ is a compact real tree and thus belongs to $\mathbb{T}$. Define a nonnegative measure Ψ_T on $\mathbb{T} \times \mathbb{R}_+$ by, for every $f \in \mathcal{B}_+(\mathbb{T} \times \mathbb{R}_+)$,

$$\Psi_T(f) = \int_T \mu(\mathrm{d}x) \int_0^{H(x)} f(T_{r,x}, r)\,\mathrm{d}r. \tag{3.1}$$

As we will consider functions depending only on the mass and height of the subtrees, we introduce the measure $\Psi_T^{\mathfrak{m}\mathfrak{h}}$ on $\mathbb{R}_+^2$ defined by, for every $f \in \mathcal{B}_+(\mathbb{R}_+^2)$,

$$\Psi_T^{\mathfrak{m}\mathfrak{h}}(f) = \int_T \mu(\mathrm{d}x) \int_0^{H(x)} f(\mathfrak{m}(T_{r,x}), \mathfrak{h}(T_{r,x}))\,\mathrm{d}r. \tag{3.2}$$

Lemma 3.1. *Let T be a compact real tree. The mapping $(r,x) \mapsto T_{r,x}$ from $\{(r,x) \in \mathbb{R}_+ \times T\colon r \leq H(x)\}$ to $\mathbb{T}$ is measurable with respect to the Borel σ-fields. Furthermore, the measure Ψ_T is finite and does not depend on the choice of representative in the equivalence class in $\mathbb{T}$ of T.*

Proof. Let $(T,\emptyset,d,\mu)$ be a compact real tree and set $A \coloneqq \{(r,x) \in \mathbb{R}_+ \times T\colon r \leq H(x)\}$. For every $(r,x) \in A$, recall that $x_r \in T$ is the unique ancestor of x with height $H(x_r) = r$. We start by showing that the mapping $(r,x) \mapsto x_r$ is continuous from A to T. Let $(r,x),(s,y) \in A$. Without loss of generality, we can assume that $r \geq s$. If $H(x \wedge y) \geq s$, then we have $y_s \preccurlyeq x$ and thus $y_s \preccurlyeq x_r$. This implies that $d(x_r,y_s) = r - s$. If $H(x \wedge y) < s$, then we have $x_r \in [\![x \wedge y, x]\!]$ and $y_s \in [\![x \wedge y, y]\!]$. This implies that x_r and y_s belong to $[\![x,y]\!]$, and thus $d(x_r,y_s) \leq d(x,y)$. In all cases, we have

$$d(x_r,y_s) \leq d(x,y) + |r-s|.$$

This proves that $(r,x) \mapsto x_r$ is continuous.

The mapping $y \mapsto T_y$ from T to $\mathbb{T}$ is continuous from below, in the sense that for $y \in T$

$$\lim_{\substack{z \to y \\ z \preccurlyeq y}} d_{\mathrm{GHP}}(T_z, T_y) = 0. \tag{3.3}$$

To see this, let $\delta > 0$, $y \in T$ and $(y_n,\, n \in \mathbb{N})$ be a sequence in T converging to y such that $y_n \preccurlyeq y$ for every $n \in \mathbb{N}$. Notice that since T is compact, it holds that there is a finite number of subtrees with height larger than δ attached to the branch $[\![\emptyset, y]\!]$. Thus, there are no subtrees with height larger than δ attached to $[\![y_n, y[\![$ for n larger than some n_0. Moreover, since $T_y = \bigcap_{n\in\mathbb{N}} T_{y_n}$, we get that $\lim_{n\to\infty} \mu(T_{y_n}) = \mu(T_y)$ implying that the mass of the subtrees attached to $[\![y_n, y[\![$ goes to 0 as n goes to infinity.

Define a correspondence between T_{y_n} and T_y by

$$\mathcal{R} \coloneqq \{(z,z) \colon\ z \in T_y\} \bigcup \{(z,y) \colon\ z \in T_{y_n} \setminus T_y\}.$$

It is straightforward to check that $\mathrm{dis}(\mathcal{R}) \leq 2(\delta + d(y_n, y))$ for $n \geq n_0$. Consider the measure on $T_{y_n} \times T_y$ defined by $m(\mathrm{d}x, \mathrm{d}z) = \mu_{|T_y}(\mathrm{d}z)\delta_z(\mathrm{d}x) = \mu_{|T_y}(\mathrm{d}x)\delta_x(\mathrm{d}z)$. Then we have $\mathrm{D}(m; \mu_{|T_{y_n}}, \mu_{|T_y}) \leq \mu(T_{y_n}) - \mu(T_y)$ and $m(\mathcal{R}^c) = 0$. It follows from (2.2) that

$$\limsup_{n\to\infty} d_{\mathrm{GHP}}(T_{y_n}, T_y) \leq \limsup_{n\to\infty} \left(\delta + d(y_n, y) + \mu(T_{y_n}) - \mu(T_y)\right) = \delta.$$

Since $\delta > 0$ is arbitrary, (3.3) readily follows.

Now it is not difficult to see that the continuity from below (3.3) of the mapping $y \mapsto T_y$ implies its measurability. By composition, it follows that the mapping $(r,x) \mapsto T_{r,x} = T_{x_r}$ from A to $\mathbb{T}$ is measurable.

Next, notice that Ψ_T is finite since

$$\Psi_T(1) = \int_T H(x)\mu(\mathrm{d}x) \leq \mathfrak{h}(T)\mathfrak{m}(T) < \infty.$$

Finally, let $f \in \mathcal{B}_+(\mathbb{T} \times \mathbb{R}_+)$ and $(T, \emptyset, d, \mu), (T', \emptyset', d', \mu')$ be two compact real trees such that there is a measure-preserving and root-preserving isometry $\varphi \colon T \to T'$. This means that φ is an isometry satisfying $\mu' = \mu \circ \varphi^{-1}$ and $\varphi(\emptyset) = \emptyset'$. Moreover, for every $x, y \in T$, since $H(x \wedge y) = 2^{-1}\left(d(\emptyset, x) + d(\emptyset, y) - d(x,y)\right)$, we deduce that

$$H(x \wedge y) = H(\varphi(x) \wedge \varphi(y)).$$

Using this and the definitions of $T_{r,x}$ and $T'_{r,\varphi(x)}$, it is easy to see that, for every $x \in T$ and $r \in [0, H(x)]$, φ induces a measure-preserving and root-preserving isometry from $T_{r,x}$ to $T'_{r,\varphi(x)}$ and therefore $f(T_{r,x}, r) = f(T'_{r,\varphi(x)}, r)$. Since $H(x) = H(\varphi(x))$, it follows that

$$\begin{aligned}
\Psi_T(f) &= \int_T \mu(\mathrm{d}x) \int_0^{H(x)} f(T_{r,x}, r)\,\mathrm{d}r \\
&= \int_T \mu(\mathrm{d}x) \int_0^{H(\varphi(x))} f(T'_{r,\varphi(x)}, r)\,\mathrm{d}r \\
&= \int_{T'} \mu \circ \varphi^{-1}(\mathrm{d}y) \int_0^{H(y)} f(T'_{r,y}, r)\,\mathrm{d}r \\
&= \Psi_{T'}(f).
\end{aligned}$$

This proves that Ψ_T does not depend on the choice of representative in the equivalence class of T which completes the proof. $\square$

Recall that $\mathrm{Lf}(T)$ is the set of leaves of T. It is well known that there exists a unique σ-finite measure ℓ on $(T, \mathcal{B}(T))$, called the length measure, such that $\ell(\mathrm{Lf}(T)) = 0$ and $\ell([\![x, y]\!]) = d(x, y)$, see *e.g.* [20, Chapter 4, §4.3.5]. The next result gives an alternative expression for Ψ_T in terms of the length measure.

Proposition 3.2. *Let $(T, \emptyset, d, \mu)$ be a compact real tree. For every $f \in \mathcal{B}_+(\mathbb{T} \times \mathbb{R}_+)$, we have*

$$\Psi_T(f) = \int_T \mu(T_y) f(T_y, H(y))\, \ell(\mathrm{d}y). \tag{3.4}$$

Proof. Let $(T, \emptyset, d, \mu)$ be a compact real tree and $f \in \mathcal{B}_+(\mathbb{T} \times \mathbb{R}_+)$. Notice that $\{(x, y) \in T^2 \colon\ y \preccurlyeq x\} = \{(x, y) \in T^2 \colon\ d(\emptyset, x) = d(\emptyset, y) + d(x, y)\}$ is closed in T^2 and thus measurable. Moreover, the mapping $y \mapsto T_y$ is measurable from T to $\mathbb{T}$ by the proof of Lemma 3.1. Thus the mapping $(x, y) \mapsto \mathbf{1}_{\{y \preccurlyeq x\}} f(T_y, H(y))$ is measurable. By Fubini's theorem, it follows that

$$\begin{aligned}\int_T \mu(T_y) f\left(T_y, H(y)\right)\, \ell(\mathrm{d}y) &= \int_T \mu(\mathrm{d}x) \int_T \mathbf{1}_{\{y \preccurlyeq x\}} f\left(T_y, H(y)\right)\, \ell(\mathrm{d}y) \\ &= \int_T \mu(\mathrm{d}x) \int_{[\![\emptyset, x]\!]} f\left(T_y, H(y)\right)\, \ell(\mathrm{d}y).\end{aligned}$$

Let $x \in T$ and let $f_{\emptyset,x} \colon [0, H(x)] \to [\![\emptyset, x]\!]$ be the unique isometry such that $f_{\emptyset,x}(0) = \emptyset$ and $f_{\emptyset,x}(H(x)) = x$. Using that $\ell_{|[\![\emptyset,x]\!]} = \lambda \circ f_{\emptyset,x}^{-1}$ where λ is the Lebesgue measure on $[0, H(x)]$, we get that

$$\int_{[\![\emptyset,x]\!]} f\left(T_y, H(y)\right)\, \ell(\mathrm{d}y) = \int_0^{H(x)} f\left(T_{f_{\emptyset,x}(r)}, H(f_{\emptyset,x}(r))\right)\, \mathrm{d}r.$$

Since $f_{\emptyset,x}$ is an isometry, for every $r \in [0, H(x)]$, $f_{\emptyset,x}(r)$ is the unique ancestor of x at height r, that is x_r, and $H(f_{\emptyset,x}(r)) = r$. As $T_{f_{\emptyset,x}(r)} = T_{x_r} = T_{r,x}$ for every $r \in [0, H(x)]$, it follows that

$$\int_T \mu(T_y) f\left(T_y, H(y)\right)\, \ell(\mathrm{d}y) = \int_T \mu(\mathrm{d}x) \int_0^{H(x)} f\left(T_{r,x}, r\right)\, \mathrm{d}r.$$

This concludes the proof. $\square$

The main result of this section concerns the continuity of the mapping $\Psi \colon T \mapsto \Psi_T$.

Proposition 3.3. *The mapping $\Psi \colon T \mapsto \Psi_T$, from $\mathbb{T}$ endowed with the Gromov-Hausdorff-Prokhorov topology to $\mathcal{M}(\mathbb{T} \times \mathbb{R}_+)$ endowed with the topology of weak convergence, is well defined and continuous.*

The end of this section is devoted to the proof of Proposition 3.3. For T a compact real tree, $x \in T$, $s \in [0, +\infty]$, $r \in [0, s \wedge H(x)]$, we define the following set of elements of T such that their common ancestor with x has height in $[r, s]$

$$T_{[r,s],x} = \{y \in T \colon\ H(y \wedge x) \in [r, s]\}.$$

Recall that x_r is the ancestor of x at height r in T, and is also seen as the root of the tree $T_{r,x}$. We shall see $T_{[r,s],x}$ as a compact real tree rooted at x_r with measure $\mu_{|T_{[r,s],x}} = \mu(\cdot \cap T_{[r,s],x})$ and thus $T_{[r,s],x} \in \mathbb{T}$. Recall that $\mathfrak{m}(T_{[r,s],x}) = \mu(T_{[r,s],x})$ denotes its mass and $\mathfrak{h}(T_{[r,s],x}) = \sup\{H(y) \colon\ y \in T_{[r,s],x} \subset T\} - r$ its height. Notice in particular that $T_{[r,+\infty],x} = T_{r,x}$ for $r \in [0, H(x)]$.

We first establish an estimate for the Gromov-Hausdorff-Prokhorov distance between subtrees of two real trees in terms of the distance between the trees themselves.

Lemma 3.4. *Let T, T' be compact real trees and let $\delta > d_{\mathrm{GHP}}(T, T')$. Let $\mathcal{R}$ be a correspondence between T and T' such that $(\emptyset, \emptyset') \in \mathcal{R}$ and let m be a measure on $T \times T'$ such that*

$$\frac{1}{2}\,\mathrm{dis}(\mathcal{R}) \vee \mathrm{D}(m; \mu, \mu') \vee m(\mathcal{R}^c) \leq \delta.$$

Then for every (x, x') in $\mathcal{R}$ and every $r \geq 0$ such that $6\delta \leq r \leq H(x) \wedge H(x')$, we have

$$d_{\mathrm{GHP}}(T_{r,x}, T'_{r,x'}) \leq 8\delta + 2\mathfrak{m}\left(T_{[r-6\delta, r+6\delta], x}\right) + 2\mathfrak{h}(T_{[r-3\delta, r+6\delta], x}). \tag{3.5}$$

Proof. Similarly to x_r, we denote by x'_r the ancestor of x' at height r in T', which is also seen as the root of $T'_{r,x'}$. We shall bound $d_{\mathrm{GHP}}(T_{r,x}, T'_{r,x'})$ from above by

$$\frac{1}{2}\,\mathrm{dis}(\widetilde{\mathcal{R}}) \vee \mathrm{D}(\widetilde{m}; \tilde{\mu}, \tilde{\mu}') \vee \widetilde{m}(\widetilde{\mathcal{R}}^c)$$

where $\widetilde{\mathcal{R}}$ is a well chosen correspondence between $T_{r,x}$ and $T'_{r,x'}$ and $\widetilde{m}$ (resp. $\tilde{\mu}$, $\tilde{\mu}'$) is the restriction of the measure m (resp. μ, μ') to $T_{r,x} \times T'_{r,x'}$ (resp. $T_{r,x}$, $T'_{r,x'}$). We begin by noticing that, for every $(t, t'), (s, s') \in \mathcal{R}$, we have

$$|d(t, s) - d'(t', s')| \leq \mathrm{dis}(\mathcal{R}) \leq 2\delta. \tag{3.6}$$

In particular, taking $(s, s') = (\emptyset, \emptyset') \in \mathcal{R}$ yields

$$|H(t) - H(t')| \leq 2\delta. \tag{3.7}$$

Using this, we get that for $(t, t') \in \mathcal{R}$

$$\begin{aligned} H(t' \wedge x') &= \frac{1}{2}\left(H(t') + H(x') - d'(t', x')\right) \\ &\geq \frac{1}{2}\left(H(t) - 2\delta + H(x) - 2\delta - d(t, x) - 2\delta\right) \\ &= H(t \wedge x) - 3\delta. \end{aligned} \tag{3.8}$$

Step 1: we construct a correspondence between $T_{r,x}$ and $T'_{r,x'}$ and give an upper bound of its distortion. Let $(t, t') \in \mathcal{R}$. Assume that $H(t \wedge x) \geq r + 3\delta$. Then, we get that $t \in T_{r,x}$ and that $H(t' \wedge x') \geq r$ by (3.8), that is $t' \in T'_{r,x'}$. This gives that $(t, t') \in T_{r,x} \times T'_{r,x'}$. Similarly, if $H(t' \wedge x') \geq r + 3\delta$, we get $(t, t') \in T_{r,x} \times T'_{r,x'}$. Therefore, the following set

$$\widetilde{\mathcal{R}} = \{(t, t') \in \mathcal{R} \colon \max(H(t \wedge x), H(t' \wedge x')) \geq r + 3\delta\} \bigcup \left(T_{[r, r+3\delta], x} \times \{x'_r\}\right) \bigcup \left(\{x_r\} \times T'_{[r, r+3\delta], x'}\right)$$

is a correspondence between $T_{r,x}$ and $T'_{r,x'}$. We give a bound of its distortion. Let $(t, t'), (s, s') \in \widetilde{\mathcal{R}}$.
Case 1: Assume that $(t, t') \in \mathcal{R}$ and $(s, s') \in \mathcal{R}$, then by (3.6) we have

$$|d(t, s) - d'(t', s')| \leq 2\delta.$$

Case 2: Assume that $(t, t') \in \mathcal{R}$ and $(s, s') \notin \mathcal{R}$. Without loss of generality, we may assume that $s = x_r$ and thus $H(s' \wedge x') \in [r, r + 3\delta]$. Let $y' \in T'$ such that $(x_r, y') \in \mathcal{R}$, then using (3.6) and the triangle inequality, we get

$$\begin{aligned} |d(t, s) - d'(t', s')| &\leq |d(t, x_r) - d'(t', y')| + |d'(t', y') - d'(t', s')| \\ &\leq 2\delta + d'(y', s') \\ &\leq 2\delta + d'(y', x'_r) + d'(x'_r, s'). \end{aligned}$$

Notice that by (3.8), we have $H(y' \wedge x') \geq H(x_r \wedge x) - 3\delta = r - 3\delta$, so either $H(y' \wedge x') \geq r$ or $H(y' \wedge x') \in [r - 3\delta, r)$. In the first case, x'_r is necessarily an ancestor of y' and we have $H(y' \wedge x'_r) = r$. In the second case, we have $y' \wedge x' = y' \wedge x'_r$ and $H(y' \wedge x'_r) \geq r - 3\delta$. Thus, in all cases we have $H(y' \wedge x'_r) \geq r - 3\delta$ and then

$$d'(y', x'_r) = H(y') + H(x'_r) - 2H(y' \wedge x'_r) \leq H(x_r) + 2\delta + r - 2(r - 3\delta) = 8\delta.$$

On the other hand, since we assumed that $H(s' \wedge x') \in [r, r + 3\delta]$, we get that x'_r is an ancestor of s' and $s' \in T'_{[r,r+3\delta],\, x'}$. We deduce that

$$d'(x'_r, s') = H(s') - H(x'_r) = H(s') - r \leq \mathfrak{h}(T'_{[r,r+3\delta],\, x'}). \tag{3.9}$$

It follows that

$$|d(t,s) - d(t', s')| \leq 10\delta + \mathfrak{h}(T'_{[r,r+3\delta],\, x'}).$$

Case 3: Assume that $(t, t'), (s, s') \notin \mathcal{R}$.

Case 3a. If $t = s = x_r$, then necessarily $H(t' \wedge x'), H(s' \wedge x') \in [r, r + 3\delta)$. Arguing as in (3.9), we have

$$|d(t,s) - d'(t', s')| = d'(t', s') \leq d'(t', x'_r) + d'(x'_r, s') \leq 2\mathfrak{h}(T'_{[r,r+3\delta],\, x'}).$$

Case 3b. If $s = x_r$ and $t' = x'_r$, then by the same argument we used to get (3.9), we have

$$|d(t,s) - d(t', s')| \leq d(t, x_r) + d(x'_r, s') \leq \mathfrak{h}(T_{[r,r+3\delta],\, x}) + \mathfrak{h}(T'_{[r,r+3\delta],\, x'}).$$

It follows that

$$\mathrm{dis}(\widetilde{\mathcal{R}}) \leq 10\delta + 2\mathfrak{h}(T_{[r,r+3\delta],\, x}) + 2\mathfrak{h}(T'_{[r,r+3\delta],\, x'}). \tag{3.10}$$

Step 2: we define a measure on $T_{r,\,x} \times T'_{r,\,x'}$ and give an upper bound of its discrepancy. Denote by $\widetilde{m}$ the restriction of the measure m to $T_{r,\,x} \times T'_{r,\,x'}$. Let $A \subset T_{r,\,x}$ be a Borel set. We have $\widetilde{m} \circ \widetilde{p}^{-1}(A) = \widetilde{m}(A \times T'_{r,\,x'}) = m(A \times T'_{r,\,x'})$ where $\widetilde{p}\colon T_{r,\,x} \times T'_{r,\,x'} \to T_{r,\,x}$ is the canonical projection. Notice that

$$\begin{aligned} m\left(A \times T'\right) - m\left(A \times T'_{r,\,x'}\right) &= m\left(A \times (T' \setminus T'_{r,\,x'})\right) \\ &= m\left(A \times (T' \setminus T'_{r,\,x'}) \cap \mathcal{R}\right) + m\left(A \times (T' \setminus T'_{r,\,x'}) \cap \mathcal{R}^c\right) \\ &\leq m\left(A \times (T' \setminus T'_{r,\,x'}) \cap \mathcal{R}\right) + \delta. \end{aligned}$$

For $(t, t') \in \left(A \times (T' \setminus T'_{r,\,x'})\right) \cap \mathcal{R}$, using (3.8) and the fact that $A \subset T_{r,\,x}$, we get

$$H(t' \wedge x') \geq H(t \wedge x) - 3\delta \geq r - 3\delta.$$

Moreover, we have $H(t' \wedge x') < r < r + 3\delta$ since $t' \notin T'_{r,\,x'}$. This gives the inclusion $\left(A \times (T' \setminus T'_{r,\,x'})\right) \cap \mathcal{R} \subset T \times T'_{[r-3\delta,r+3\delta],\, x'}$. As $d_{\mathrm{TV}}(m \circ p'^{-1}, \mu') \leq \mathrm{D}(m; \mu, \mu') \leq \delta$, we deduce that

$$\begin{aligned} m\left(A \times T'\right) - m\left(A \times T'_{r,\,x'}\right) &\leq m\left(T \times T'_{[r-3\delta,r+3\delta],\, x'}\right) + \delta \\ &\leq \mu'\left(T'_{[r-3\delta,r+3\delta],\, x'}\right) + d_{\mathrm{TV}}(m \circ p'^{-1}, \mu') + \delta \\ &\leq \mu'\left(T'_{[r-3\delta,r+3\delta],\, x'}\right) + 2\delta. \end{aligned}$$

Recall that $\widetilde{\mu}$ is the restriction of the measure μ to $T_{r,x}$. It follows that

$$\begin{aligned} \left|\widetilde{m} \circ \widetilde{p}^{-1}(A) - \widetilde{\mu}(A)\right| &= \left|m\left(A \times T'_{r,\,x'}\right) - \mu(A)\right| \\ &\leq \left|m\left(A \times T'_{r,\,x'}\right) - m(A \times T')\right| + \left|m\left(A \times T'\right) - \mu(A)\right| \end{aligned}$$

$$\begin{aligned}&\leq \left|m\left(A\times T'_{r,x'}\right)-m(A\times T')\right|+\mathrm{D}(m;\mu,\mu')\\&\leq \mu'\left(T'_{[r-3\delta,r+3\delta],x'}\right)+3\delta.\end{aligned}$$

By symmetry, we deduce that

$$\mathrm{D}(\widetilde{m};\widetilde{\mu},\widetilde{\mu}')\leq \mathfrak{m}\Big(T_{[r-3\delta,r+3\delta],x}\Big)+\mathfrak{m}\Big(T'_{[r-3\delta,r+3\delta],x'}\Big)+6\delta. \tag{3.11}$$

Step 3: we give an upper bound of $\widetilde{m}(\widetilde{\mathcal{R}}^c)$. Let $(t,t')\in T_{r,x}\times T_{r,x'}\setminus\widetilde{\mathcal{R}}$. If $H(t\wedge x)>r+3\delta$ then necessarily $(t,t')\notin\mathcal{R}$ by our construction of $\widetilde{\mathcal{R}}$. Therefore, we have

$$\begin{aligned}m(\widetilde{\mathcal{R}}^c)=m(T_{r,x}\times T'_{r,x'}\setminus\widetilde{\mathcal{R}})&=m\left((t,t')\in T_{r,x}\times T'_{r,x'}\setminus\widetilde{\mathcal{R}}\colon\ H(t\wedge x)>r+3\delta\right)\\&\quad+m\left((t,t')\in T_{r,x}\times T'_{r,x'}\colon\ H(t\wedge x)\in[r,r+3\delta]\right)\\&\leq m(\mathcal{R}^c)+\mu\left(T_{[r,r+3\delta],x}\right)+d_{\mathrm{TV}}(m\circ p^{-1},\mu)\\&\leq \mathfrak{m}\Big(T_{[r,r+3\delta],x}\Big)+2\delta.\end{aligned}\tag{3.12}$$

Step 4: we can now conclude. Combining (3.10), (3.11) and (3.12) and using the definition of the Gromov-Hausdorff-Prokhorov distance, we get

$$d_{\mathrm{GHP}}(T_{r,x},T'_{r,x'})\leq 6\delta+\mathfrak{m}\Big(T_{[r-3\delta,r+3\delta],x}\Big)+\mathfrak{m}\Big(T'_{[r-3\delta,r+3\delta],x'}\Big)+\mathfrak{h}\Big(T_{[r,r+3\delta],x}\Big)+\mathfrak{h}\Big(T'_{[r,r+3\delta],x'}\Big).\tag{3.13}$$

First, notice that

$$\begin{aligned}\mathfrak{m}\Big(T'_{[r-3\delta,r+3\delta],x'}\Big)&=\mu'\left(t'\colon\ H(t'\wedge x')\in[r-3\delta,r+3\delta]\right)\\&\leq m\left((t,t')\colon\ H(t'\wedge x')\in[r-3\delta,r+3\delta]\right)+d_{\mathrm{TV}}(m\circ {p'}^{-1},\mu')\\&\leq m\left((t,t')\in\mathcal{R}\colon\ H(t'\wedge x')\in[r-3\delta,r+3\delta]\right)+m\left(\mathcal{R}^c\right)+\delta\\&\leq m\left((t,t')\in\mathcal{R}\colon\ H(t'\wedge x')\in[r-3\delta,r+3\delta]\right)+2\delta.\end{aligned}$$

Using (3.8), we get by symmetry that, for $(t,t')\in\mathcal{R}$,

$$H(t'\wedge x')-3\delta\leq H(t\wedge x)\leq H(t'\wedge x')+3\delta.\tag{3.14}$$

We deduce that

$$\begin{aligned}\mathfrak{m}\Big(T'_{[r-3\delta,r+3\delta],x'}\Big)&\leq m\left((t,t')\in\mathcal{R}\colon\ H(t\wedge x)\in[r-6\delta,r+6\delta]\right)+2\delta\\&\leq m\left((t,t')\colon\ H(t\wedge x)\in[r-6\delta,r+6\delta]\right)+2\delta\\&\leq \mu\left(t\colon\ H(t\wedge x)\in[r-6\delta,r+6\delta]\right)+d_{\mathrm{TV}}(m\circ p^{-1},\mu)+2\delta\\&\leq \mathfrak{m}\Big(T_{[r-6\delta,r+6\delta],x}\Big)+3\delta.\end{aligned}\tag{3.15}$$

Secondly, let $t'\in T'_{[r,r+3\delta],x'}$. We have $H(t'\wedge x')\in[r,r+3\delta]$. Let $t\in T$ such that $(t,t')\in\mathcal{R}$. Thanks to (3.14), we get $H(t\wedge x)\in[r-3\delta,r+6\delta]$. Since $(t,t')\in\mathcal{R}$, we also have $|H(t')-H(t)|\leq 2\delta$ by(3.7). We deduce that

$$\begin{aligned}\mathfrak{h}\Big(T'_{[r,r+3\delta],x'}\Big)&=\sup\left\{H(t')\colon\ t'\in T',H(t'\wedge x')\in[r,r+3\delta]\right\}-r\\&\leq \sup\left\{H(t)\colon\ t\in T,H(t\wedge x)\in[r-3\delta,r+6\delta]\right\}-r+2\delta\\&=\mathfrak{h}\Big(T_{[r-3\delta,r+6\delta],x}\Big)-\delta.\end{aligned}\tag{3.16}$$

Using (3.15) and (3.16) in conjunction with (3.13) yields the result. □

Proof of Proposition 3.3. Fix a compact real tree $T = (T, d, \emptyset, \mu)$. We will show that $\Psi_{T'} \to \Psi_T$ weakly as $T' \to T$ for d_{GHP}. Let $\varepsilon > 0$ and let $T' = (T', d', \emptyset', \mu')$ be a compact real tree such that $d_{\mathrm{GHP}}(T, T') \leq \varepsilon$. Then there exist a correspondence $\mathcal{R}$ between T and T' and a measure m on $T \times T'$ such that $(\emptyset, \emptyset') \in \mathcal{R}$, $m(\mathcal{R}^c) \leq \varepsilon$, $\mathrm{dis}(\mathcal{R}) \leq 2\varepsilon$ and $\mathrm{D}(m; \mu, \mu') \leq \varepsilon$. In particular, we will make constant use of the inequalities $|m(T \times T') - \mathfrak{m}(T)| \leq \varepsilon$ and $|H(x) - H(x')| \leq 2\varepsilon$ for $(x, x') \in \mathcal{R}$. Let $f \in \mathcal{C}_b(\mathbb{T} \times \mathbb{R}_+)$ be Lipschitz. Write

$$\Psi_T(f) - \Psi_{T'}(f) = A_1 + A_2 + A_3 + A_4,$$

where

$$\begin{aligned}
A_1 &= \int_T \mu(\mathrm{d}x) \int_0^{H(x)} f(T_{r,x}, r)\mathrm{d}r - \int_T m \circ p^{-1}(\mathrm{d}x) \int_0^{H(x)} f(T_{r,x}, r)\mathrm{d}r \\
A_2 &= \int_{\mathcal{R}} m(\mathrm{d}x, \mathrm{d}x') \left(\int_0^{H(x)} f(T_{r,x}, r)\mathrm{d}r - \int_0^{H(x')} f(T'_{r,x'}, r)\mathrm{d}r \right) \\
A_3 &= \int_{\mathcal{R}^c} m(\mathrm{d}x, \mathrm{d}x') \left(\int_0^{H(x)} f(T_{r,x}, r)\mathrm{d}r - \int_0^{H(x')} f(T'_{r,x'}, r)\mathrm{d}r \right) \\
A_4 &= \int_{T'} m \circ p'^{-1}(\mathrm{d}x') \int_0^{H(x')} f(T'_{r,x'}, r)\mathrm{d}r - \int_{T'} \mu(\mathrm{d}x') \int_0^{H(x')} f(T'_{r,x'}, r)\mathrm{d}r.
\end{aligned}$$

Notice that

$$|A_1| \leq 2 d_{\mathrm{TV}}(m \circ p^{-1}, \mu) \sup_{x \in T} \int_0^{H(x)} f(T_{r,x}, r)\mathrm{d}r \leq 2\mathfrak{h}(T) \|f\|_\infty \varepsilon. \tag{3.17}$$

Similarly, we have

$$|A_4| \leq 2\mathfrak{h}(T') \|f\|_\infty \varepsilon \leq 2(\mathfrak{h}(T) + 2\varepsilon) \|f\|_\infty \varepsilon, \tag{3.18}$$

where in the second inequality we used that $\mathfrak{h}(T') \leq \mathfrak{h}(T) + 2d_{\mathrm{GHP}}(T, T') \leq \mathfrak{h}(T) + 2\varepsilon$ by (2.3). Next, we have

$$|A_3| \leq m(\mathcal{R}^c)(\mathfrak{h}(T) + \mathfrak{h}(T')) \|f\|_\infty \leq 2(\mathfrak{h}(T) + \varepsilon) \|f\|_\infty \varepsilon. \tag{3.19}$$

We now provide a bound for A_2. We have

$$\begin{aligned}
A_2 = &\int_{\mathcal{R}} \mathbf{1}_{\{H(x) \geq H(x')\}} m(\mathrm{d}x, \mathrm{d}x') \left(\int_0^{H(x)} f(T_{r,x}, r)\mathrm{d}r - \int_0^{H(x')} f(T'_{r,x'}, r)\mathrm{d}r \right) \\
&+ \int_{\mathcal{R}} \mathbf{1}_{\{H(x) < H(x')\}} m(\mathrm{d}x, \mathrm{d}x') \left(\int_0^{H(x)} f(T_{r,x}, r)\mathrm{d}r - \int_0^{H(x')} f(T'_{r,x'}, r)\mathrm{d}r \right).
\end{aligned} \tag{3.20}$$

We only treat the first term, the second one being similar. We have

$$\begin{aligned}
&\int_{\mathcal{R}} \mathbf{1}_{\{H(x) \geq H(x')\}} m(\mathrm{d}x, \mathrm{d}x') \left(\int_0^{H(x)} f(T_{r,x}, r)\mathrm{d}r - \int_0^{H(x')} f(T'_{r,x'}, r)\mathrm{d}r \right) \\
&\quad = \int_{\mathcal{R}} \mathbf{1}_{\{H(x) \geq H(x')\}} m(\mathrm{d}x, \mathrm{d}x') \left(\int_0^{H(x')} \left(f(T_{r,x}, r) - f(T'_{r,x'}, r) \right) \mathrm{d}r + \int_{H(x')}^{H(x)} f(T_{r,x}, r)\mathrm{d}r \right).
\end{aligned}$$

On the one hand, we get

$$\left| \int_{\mathcal{R}} \mathbf{1}_{\{H(x) \geq H(x')\}} m(\mathrm{d}x, \mathrm{d}x') \int_{H(x')}^{H(x)} f(T_{r,x}, r)\mathrm{d}r \right| \leq \int_{\mathcal{R}} \|f\|_\infty |H(x) - H(x')| m(\mathrm{d}x, \mathrm{d}x')$$

$$\begin{aligned}&\leq \|f\|_\infty \, m(T\times T')\,\mathrm{dis}(\mathcal{R})\\ &\leq 2\,\|f\|_\infty\,(\mathfrak{m}(T)+\varepsilon)\varepsilon. \end{aligned}\tag{3.21}$$

On the other hand, we have

$$\begin{aligned}&\left|\int_{\mathcal{R}} \mathbf{1}_{\{H(x)\geq H(x')\}} m(\mathrm{d}x,\mathrm{d}x')\int_0^{H(x')}\Big(f(T_{r,x},r)-f(T'_{r,x},r)\Big)\,\mathrm{d}r\right|\\
&\leq \|f\|_{\mathrm{L}}\int_{\mathcal{R}} \mathbf{1}_{\{H(x)\geq H(x')\}} m(\mathrm{d}x,\mathrm{d}x')\int_0^{H(x')} d_{\mathrm{GHP}}\left(T_{r,x},T'_{r,x}\right)\mathbf{1}_{\{r\geq 6\varepsilon\}}\mathrm{d}r+\int_{\mathcal{R}} m(\mathrm{d}x,\mathrm{d}x')\int_0^{6\varepsilon}2\,\|f\|_\infty\,\mathrm{d}r\\
&\leq 2\,\|f\|_{\mathrm{L}}\int m(\mathrm{d}x,\mathrm{d}x')\int_0^{H(x)}\Big(\mathfrak{m}(T_{[r-3\varepsilon,r+6\varepsilon],x})+\mathfrak{h}(T_{[r-6\varepsilon,r+6\varepsilon],x})\Big)\,\mathbf{1}_{\{r\geq 6\varepsilon\}}\,\mathrm{d}r\\
&\qquad\qquad + 8\,\|f\|_{\mathrm{L}}\,\mathfrak{h}(T)(\mathfrak{m}(T)+\varepsilon)\varepsilon+12\,\|f\|_\infty\,(\mathfrak{m}(T)+\varepsilon)\varepsilon. \end{aligned}\tag{3.22}$$

where we used (3.5) for the last inequality. Using Fubini's theorem, we get

$$\begin{aligned}&\int m(\mathrm{d}x,\mathrm{d}x')\int_0^{H(x)}\mathfrak{m}(T_{[r-6\varepsilon,r+6\varepsilon],x})\mathbf{1}_{\{r\geq 6\varepsilon\}}\,\mathrm{d}r\\
&\qquad\qquad =\int m(\mathrm{d}x,\mathrm{d}x')\int_0^{H(x)}\mu(t\colon\ H(t\wedge x)\in[r-6\varepsilon,r+6\varepsilon])\mathbf{1}_{\{r\geq 6\varepsilon\}}\,\mathrm{d}r\\
&\qquad\qquad =\int m(\mathrm{d}x,\mathrm{d}x')\int_T\mu(\mathrm{d}t)\int_0^{H(x)}\mathbf{1}_{\{H(t\wedge x)\in[r-6\varepsilon,r+6\varepsilon]\}}\mathbf{1}_{\{r\geq 6\varepsilon\}}\,\mathrm{d}r\\
&\qquad\qquad \leq 12\,\mathfrak{m}(T)(\mathfrak{m}(T)+\varepsilon)\varepsilon. \end{aligned}\tag{3.23}$$

Moreover, since T is compact, it holds that for every $x\in T$ and every $\delta>0$, there is a finite number of subtrees with height larger than δ attached to the branch $[\![\emptyset,x]\!]$. Let $r\in(0,H(x))$. Recall that x_r is the unique ancestor of x with height $H(x_r)=r$. Assume that x_r is not a branching point. Then, for every $\delta>0$ and for $\varepsilon>0$ small enough (depending on δ), there are no subtrees with height larger than δ attached to $[\![x_{r-3\varepsilon},x_{r+6\varepsilon}]\!]$. (To be precise, if $y\in[\![x_{r-3\varepsilon},x_{r+6\varepsilon}]\!]$ is a branching point, the tree attached at y is $T_{[s,s],x}$ with $s=H(y)$). Therefore, we have $\mathfrak{h}(T_{[r-3\varepsilon,r+6\varepsilon],x})\leq\delta+9\varepsilon$. This proves that, for every $r\in(0,H(x))$ such that x_r is not a branching point,

$$\lim_{\varepsilon\to 0}\mathfrak{h}(T_{[r-3\varepsilon,r+6\varepsilon],x})=0.\tag{3.24}$$

But since T is compact, there are (at most) countably many $r\in(0,H(x))$ such that x_r is a branching point. It follows that (3.24) holds for every $x\in T$ and $\mathrm{d}r$-a.e. $r\in[0,H(x)]$. Notice that $\mathfrak{h}(T_{r-3\varepsilon,r+6\varepsilon,x})\leq\mathfrak{h}(T)$ and the measure $\mathbf{1}_{\{0\leq r\leq H(x)\}}\,\mu(\mathrm{d}x)\mathrm{d}r$ is finite as its total mass is less than $\mathfrak{h}(T)\mathfrak{m}(T)$ which is finite. We get by the dominated convergence theorem that

$$\lim_{\varepsilon\to 0}\int_T\mu(\mathrm{d}x)\int_0^{H(x)}\mathfrak{h}(T_{[r-3\varepsilon,r+6\varepsilon],x})\mathbf{1}_{\{r\geq 6\varepsilon\}}\,\mathrm{d}r=0.$$

Since

$$\left|\int_T\Big(m\circ p^{-1}(\mathrm{d}x)-\mu(\mathrm{d}x)\Big)\int_0^{H(x)}\mathfrak{h}(T_{[r-3\varepsilon,r+6\varepsilon],x})\mathbf{1}_{\{r\geq 6\varepsilon\}}\,\mathrm{d}r\right|\leq 2\mathfrak{h}(T)^2 d_{\mathrm{TV}}(m\circ p^{-1},\mu)\leq 2\mathfrak{h}(T)^2\varepsilon,$$

it follows that

$$\lim_{\varepsilon\to 0}\int m(\mathrm{d}x,\mathrm{d}x')\int_0^{H(x)}\mathfrak{h}(T_{[r-3\varepsilon,r+6\varepsilon],x})\mathbf{1}_{\{r\geq 6\varepsilon\}}\,\mathrm{d}r=0.\tag{3.25}$$

Thus, by (3.17)–(3.19), (3.21)–(3.23) and (3.25), we deduce that

$$\lim_{\varepsilon\to 0}\sup_{d_{\mathrm{GHP}}(T,T')<\varepsilon}\Psi_{T'}(f)=\Psi_T(f)$$

for every Lipschitz function $f \in \mathcal{C}_b(\mathbb{T} \times \mathbb{R}_+)$. This proves that $\Psi \colon \mathbb{T} \to \mathcal{M}(\mathbb{T} \times \mathbb{R}_+)$ is continuous which concludes the proof. $\square$

4. Bienaymé-Galton-Watson trees and stable Lévy trees

Throughout this work, we fix a random variable ξ whose distribution is critical and belongs to the domain of attraction of a stable distribution with index $\gamma \in (1, 2]$. More precisely, we assume that ξ takes values in $\mathbb{N} = \{0, 1, 2, \ldots\}$ and that it satisfies the following conditions:

(ξ1) ξ is critical, *i.e.* $\mathbb{E}[\xi] = 1$, and nondegenerate, *i.e.* $\mathbb{P}(\xi = 0) > 0$,

(ξ2) ξ belongs to the domain of attraction of a stable distribution with index $\gamma \in (1, 2]$, *i.e.* $\mathbb{E}\left[\xi^2 \mathbf{1}_{\{\xi \le n\}}\right] = n^{2-\gamma} L(n)$, where $L \colon \mathbb{R}_+ \to \mathbb{R}_+$ is a slowly varying function.

By [22, Theorem XVII.5.2] or [28, Theorem 5.2], assumption (ξ2) is equivalent to the existence of a positive sequence $(b_n,\, n \ge 1)$ such that, if $(\xi_n,\, n \ge 1)$ is a sequence of independent random variables with the same distribution as ξ, then

$$\frac{1}{b_n}\left(\sum_{k=1}^{n} \xi_k - n\right) \xrightarrow[n\to\infty]{(d)} X_1, \tag{4.1}$$

where $(X_t,\, t \ge 0)$ is a strictly stable spectrally positive Lévy process with Laplace transform $\mathbb{E}[\exp(-\lambda X_t)] = \exp(t\kappa\lambda^\gamma)$ where $\gamma \in (1, 2]$ and $\kappa > 0$. Note that we have automatically $b_n/n \to 0$ as $n \to \infty$. In most of our results, we make the following stronger assumption on ξ:

(ξ2)′ $\mathbb{E}\left[\xi^2 \mathbf{1}_{\{\xi \le n\}}\right] = n^{2-\gamma} L(n)$ where $L \colon \mathbb{R}_+ \to \mathbb{R}_+$ is a slowly varying function which is bounded away from zero and infinity.

Assumption (ξ2)′ is equivalent to the normalizing sequence $(b_n,\, n \ge 1)$ which appears in (4.1) satisfying

$$\underline{b} n^{1/\gamma} \le b_n \le \overline{b} n^{1/\gamma}, \quad \forall n \ge 1, \tag{4.2}$$

for some constants $0 < \underline{b} < \overline{b} < \infty$. Indeed, if $\gamma = 2$, we have the convergence of $n b_n^{-2} L(b_n)$ to some positive constant by [28, Theorem 5.2 and Eq. (5.44)]. Similarly, if $\gamma \in (1, 2)$, using [28, Theorem 5.3 and Eq. (5.7)], we have as $n \to \infty$ that

$$n\,\mathbb{P}(\xi > b_n) \sim \frac{2-\gamma}{\gamma} n b_n^{-\gamma} L(b_n).$$

On the other hand, [28, Eq. (5.10)] entails the convergence of $n\,\mathbb{P}(\xi > b_n)$ to some positive constant. Therefore, for $\gamma \in (1, 2]$, the sequence $n^{1/\gamma} b_n^{-1} L(b_n)^{1/\gamma}$ converges to some positive constant. Thus, if L is bounded away from 0 and infinity, then (4.2) follows. The proof of the converse (which we shall not use) is left for the reader.

4.1. Results on conditioned Bienaymé-Galton-Watson trees.

Results on conditioned Bienaymé-Galton-Watson trees. Recall that the span of the integer-valued random variable ξ is the largest integer λ_0 such that a.s. $\xi \in a + \lambda_0\mathbb{Z}$ for some $a \in \mathbb{Z}$. As we only consider ξ with $\mathbb{P}(\xi = 0) > 0$, the span is the largest integer λ_0 such that a.s. $\xi \in \lambda_0\mathbb{Z}$, *i.e.* the greatest common divisor of $\{k \ge 1 \colon \ \mathbb{P}(\xi = k) > 0\}$.

Assume that ξ satisfies (ξ1) and (ξ2) and denote by $\mathfrak{g}$ the density of the random variable X_1 appearing in (4.1). Then the function $\mathfrak{g}$ is continuous on $\mathbb{R}$ (in fact infinitely differentiable) and satisfies

$$\mathfrak{g}(0) = \frac{1}{\kappa^{1/\gamma}\,|\Gamma(-1/\gamma)|}, \tag{4.3}$$

where Γ is Euler's gamma function, see [22, Lemma XVII.6.1] or [28, Example 3.15 and Eq. (4.6)]. In particular, when $\gamma = 2$, $\mathfrak{g}$ is the density of a centered Gaussian distribution with variance 2κ and we have

$$\mathfrak{g}(0) = \frac{1}{2\sqrt{\kappa\pi}}\cdot \tag{4.4}$$

Recall that $(\xi_n,\, n \geq 1)$ is a sequence of independent random variables with the same distribution as ξ and define $S_n = \sum_{k=1}^n \xi_k$. The following result is a direct consequence of the local limit theorem, see *e.g.* [26, Chapter 4, Theorem 4.2.1].

Lemma 4.1 (Local limit theorem)**.** *Assume that ξ satisfies (ξ1) and (ξ2) and denote its span by λ_0. We have*

$$\lim_{n\to\infty} \sup_{k\geq 0} \left| \frac{b_n}{\lambda_0}\, \mathbb{P}\left(S_n = \lambda_0 k\right) - \mathfrak{g}\left(\frac{\lambda_0 k - n}{b_n}\right)\right| = 0,$$

where $\mathfrak{g}$ is the density of the random variable X_1 defined in (4.1)*. In particular, for any fixed $k \geq 0$, we have as $n \to \infty$ with $n \equiv k \,(\mathrm{mod}\ \lambda_0)$,*

$$\mathbb{P}\left(S_n = n - k\right) \sim \frac{\lambda_0 \mathfrak{g}(0)}{b_n}\cdot \tag{4.5}$$

Let τ be a BGW(ξ) tree, see *e.g.* Athreya and Ney [10]. By the well-known Otter-Dwass formula, we have, for every $n \geq 1$,

$$\mathbb{P}\left(|\tau| = n\right) = \frac{1}{n}\, \mathbb{P}\left(S_n = n - 1\right). \tag{4.6}$$

In particular, we get $\mathbb{P}\left(|\tau| = n\right) = 0$ if $n \not\equiv 1 \pmod{\lambda_0}$ while $\mathbb{P}\left(|\tau| = n\right) > 0$ for all large n with $n \equiv 1 \pmod{\lambda_0}$ by Lemma 4.1. We denote by Δ the support of the random variable $|\tau|$ when τ is not reduced to the root, that is

$$\Delta = \left\{n \geq 2 \colon\ \mathbb{P}\left(|\tau| = n\right) > 0\right\}. \tag{4.7}$$

In particular, the previous discussion implies that $\Delta \subset 1 + \lambda_0\mathbb{N}$ and conversely, $1 + \lambda_0 n \in \Delta$ for all large n. In what follows, we only consider $n \in \Delta$ and convergences should be understood along the set Δ.

We will also need the following sub-exponential tail bounds for the height of conditioned BGW trees, see [34, Theorem 2] and the discussion thereafter. For every $n \in \Delta$, τ^n will denote a BGW(ξ) tree conditioned to have n vertices, that is τ^n is distributed as τ conditionally on $\{|\tau| = n\}$.

Lemma 4.2. *Assume that ξ satisfies (ξ1) and (ξ2). For every $\alpha \in (0, \gamma/(\gamma-1))$ and every $\beta \in (0, \gamma)$, there exist two finite constants $C_0, c_0 > 0$ such that for every $y \geq 0$ and $n \in \Delta$, we have*

$$\mathbb{P}\left(\frac{b_n}{n}\mathfrak{h}(\tau^n) \leq y\right) \leq C_0 \exp\left(-c_0 y^{-\alpha}\right), \tag{4.8}$$

$$\mathbb{P}\left(\frac{b_n}{n}\mathfrak{h}(\tau^n) \geq y\right) \leq C_0 \exp\left(-c_0 y^{\beta}\right). \tag{4.9}$$

Remark 4.3.
(i) If moreover ξ satisfies $(\xi 2)'$, then we can take $\alpha = \gamma/(\gamma-1)$ in (4.8), see Appendix B.
(ii) If ξ has finite variance $\sigma_\xi^2 \in (0,\infty)$ (in which case $(\xi 2)'$ is satisfied), we have $\gamma = 2$ and we can take $b_n = \sigma_\xi \sqrt{n}$ in (4.1) with $\kappa = 1/2$ (this is just the central limit theorem). Then both (4.8) and (4.9) hold with $\alpha = \beta = 2$, see [5, Theorem 1.1 and Theorem 1.2].

An immediate consequence of Lemma 4.2 is the following estimate for the moments of $\mathfrak{h}(\tau^n)$ which extends [5, Corollary 1.3].

Lemma 4.4. *Assume that ξ satisfies $(\xi 1)$ and $(\xi 2)$. For every $p \in \mathbb{R}$, we have*

$$\sup_{n\in\Delta} \mathbb{E}\left[\left(\frac{b_n}{n}\mathfrak{h}(\tau^n)\right)^p\right] < \infty.$$

Proof. Let $p > 0$. Fix $\beta \in (0,\gamma)$. By Lemma 4.2, we have for every $n \in \Delta$

$$\mathbb{E}\left[\left(\frac{b_n}{n}\mathfrak{h}(\tau^n)\right)^p\right] = p\int_0^\infty y^{p-1}\,\mathbb{P}\left(\frac{b_n}{n}\mathfrak{h}(\tau^n) > y\right)\,\mathrm{d}y \le C_0 p\int_0^\infty y^{p-1}\mathrm{e}^{-c_0 y^\beta}\,\mathrm{d}y < \infty.$$

Similarly, fix $\alpha \in (0, \gamma/(\gamma-1))$ and apply Lemma 4.2 to get

$$\mathbb{E}\left[\left(\frac{b_n}{n}\mathfrak{h}(\tau^n)\right)^{-p}\right] = p\int_0^\infty y^{p-1}\,\mathbb{P}\left(\frac{b_n}{n}\mathfrak{h}(\tau^n) < \frac{1}{y}\right)\,\mathrm{d}y \le C_0 p\int_0^\infty y^{p-1}e^{-c_0 y^\alpha}\,\mathrm{d}y < \infty.$$

This proves the result. □

We end this section with the following lemma used in the proof of Remark 1.2-(vi).

Lemma 4.5. *Assume that ξ has finite variance $\sigma_\xi^2 \in (0,\infty)$. Let $\alpha', \beta \in \mathbb{R}$ such that $2\alpha' + \beta < 0$ and set $f_{\alpha',\beta}(\mathbf{t}) = |\mathbf{t}|^{\alpha'}\,\mathfrak{h}(\mathbf{t})^\beta \mathbf{1}_{\{|\mathbf{t}|>1\}}$. Then we have*

$$\mathbb{E}\left[f_{\alpha',\beta}(\tau)\right] < \infty, \quad \lim_{n\to\infty}\mathbb{E}\left[f_{\alpha',\beta}(\tau^n)^2\right] = 0 \quad \textit{and} \quad \sum_{n\in\Delta}\frac{\sqrt{\mathbb{E}\left[f_{\alpha',\beta}(\tau^n)^2\right]}}{n} < \infty.$$

Proof. We have

$$\mathbb{E}\left[f_{\alpha',\beta}(\tau)\right] = \sum_{n\in\Delta} n^{\alpha'}\,\mathbb{E}\left[\mathfrak{h}(\tau^n)^\beta\right]\mathbb{P}\left(|\tau| = n\right).$$

Using (4.6) and (4.5), (4.4) with $b_n = \sigma_\xi\sqrt{n}$, we have as $n \to \infty$ that

$$\mathbb{P}\left(|\tau| = n\right) \sim \frac{\lambda_0}{\sqrt{2\pi\sigma_\xi^2}}n^{-3/2}.$$

Since $\mathbb{E}\left[\mathfrak{h}(\tau^n)^\beta\right] = O(n^{\beta/2})$ as $n \to \infty$ by Lemma 4.4, we get that

$$\mathbb{E}\left[f_{\alpha',\beta}(\tau)\right] \le C\sum_{n\in\Delta} n^{-3/2+\alpha'+\beta/2} < \infty.$$

Applying Lemma 4.4 again gives $\mathbb{E}\left[f_{\alpha',\beta}(\tau^n)^2\right] = n^{2\alpha'}\,\mathbb{E}\left[\mathfrak{h}(\tau^n)^{2\beta}\right]\mathbf{1}_{\{n>1\}} \le M n^{2\alpha'+\beta}$ for some finite constant $M>0$, and the last term converges to 0 as $n\to\infty$. Finally, we have

$$\sum_{n\in\Delta} \frac{\sqrt{\mathbb{E}\left[f_{\alpha',\beta}(\tau^n)^2\right]}}{n} \le \sqrt{M}\sum_{n\in\Delta} n^{-1+\alpha'+\beta/2} < \infty.$$

□

4.2. **Stable Lévy trees.** Let us briefly recall the definition of the height process and the associated Lévy tree, see *e.g.* [15, 16, 33, 35]. Recall that $(X_t,\, t\ge 0)$ is a strictly stable Lévy process with Laplace exponent $\psi(\lambda) = \kappa\lambda^\gamma$ where $\gamma\in(1,2]$ and $\kappa>0$. For $\gamma\in(1,2)$, denote by π the associated Lévy measure

$$\pi(\mathrm{d}x) = \frac{\kappa\gamma(\gamma-1)}{\Gamma(2-\gamma)}\frac{\mathrm{d}x}{x^{1+\gamma}}. \tag{4.10}$$

Le Gall and Le Jan [35] proved that there exists a continuous process $(H(t),\, t\ge 0)$ called the ψ-height process such that for every $t\ge 0$, we have the following convergence in probability

$$H(t) = \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_0^t \mathbf{1}_{\{X_s < I_t^s+\varepsilon\}}\,\mathrm{d}s,$$

where $I_t^s = \inf_{[s,t]} X$. In the Brownian case, H is a (scaled) reflected Brownian motion. Let $\mathbf{N}$ be the excursion measure of H above 0 and set

$$\sigma = \inf\{s>0\colon\ H(s)=0\} \quad \text{and} \quad \mathfrak{h} = \sup_{s\ge 0} H(s) \tag{4.11}$$

for the duration of the excursion and its maximum. We choose to normalize the excursion measure $\mathbf{N}$ such that the distribution of σ under $\mathbf{N}$ is π_* given by

$$\pi_*(\mathrm{d}x) = \mathbf{N}\left[\sigma\in\mathrm{d}x\right] = \mathfrak{g}(0)\frac{\mathrm{d}x}{x^{1+1/\gamma}}, \tag{4.12}$$

with $\mathfrak{g}(0)$ given in (4.3). Furthermore, by [17, Eq. (14)], the distribution of $\mathfrak{h}$ under $\mathbf{N}$ is given by

$$\mathbf{N}\left[\mathfrak{h}>x\right] = (\kappa(\gamma-1)x)^{-1/(\gamma-1)}. \tag{4.13}$$

We have the following equality in "distribution" for the height process, see *e.g.* [18, Eq. (40)],

$$(H(xt),\, t\ge 0) \quad \text{under } x^{1/\gamma}\,\mathbf{N} \quad \overset{(d)}{=} \quad x^{1-1/\gamma}H \quad \text{under } \mathbf{N}\,.$$

Using this, one can make sense of the conditional probability measure $\mathbf{N}^{(x)}[\bullet] = \mathbf{N}[\bullet|\sigma=x]$ such that $\mathbf{N}^{(x)}$-a.s., $\sigma=x$ and

$$\mathbf{N}[\bullet] = \int_0^\infty \mathbf{N}^{(x)}[\bullet]\,\pi_*(\mathrm{d}x).$$

Informally, $\mathbf{N}^{(x)}$ can be seen as the distribution of the excursion of H with duration x. Moreover, the height process H has the following scaling property

$$(H(s),\, s\in[0,x]) \ \text{under}\ \mathbf{N}^{(x)} \quad \overset{(d)}{=} \quad \left(x^{1-1/\gamma}H(s/x),\, s\in[0,1]\right) \ \text{under}\ \mathbf{N}^{(1)}. \tag{4.14}$$

See also Lemma 6.11 for the scaling property of H and related processes.

We call the stable Lévy tree with branching mechanism $\psi(\lambda) = \kappa\lambda^\gamma$, the compact real tree $\mathcal{T}$ coded by the ψ-height process H under $\mathbf{N}^{(1)}$. See Remark 2.1 for the coding of real trees by excursion paths.

Remark 4.6. Notice that $\sigma = \mathfrak{m}(\mathcal{T}_H)$ and $\mathfrak{h} = \mathfrak{h}(\mathcal{T}_H)$ are the mass and the height of the tree $\mathcal{T}_H$ coded by the height process H under $\mathbf{N}$. Furthermore, for $s \in [0, \sigma]$, the notation $H(s)$ is consistent with the one introduced in Section 2.3 since $H(s)$ is the height of s in the tree coded by H under $\mathbf{N}$.

4.3. **Convergence of continuous functionals.** For every $n \in \Delta$, we let τ^n be a BGW(ξ) tree conditioned to have n vertices, and let $\mathcal{T}^n = (b_n/n)\tau^n$ be the associated real tree rescaled so that all edges have length b_n/n. Duquesne [15] (see also [33]) showed that the convergence in distribution

$$\mathcal{T}^n \xrightarrow[n\to\infty]{(d)} \mathcal{T} \tag{4.15}$$

holds in the space $\mathbb{T}$ where $\mathcal{T}$ is the stable Lévy tree with branching mechanism $\psi(\lambda) = \kappa\lambda^\gamma$.

The following result is an immediate consequence of Proposition 3.3. Recall from (3.1) and (3.2) the definitions of the measures Ψ_T and $\Psi_T^{\mathfrak{m}\mathfrak{h}}$.

Corollary 4.7. *Assume that ξ satisfies (ξ1) and (ξ2). Let τ^n be a BGW(ξ) tree conditioned to have n vertices and let $\mathcal{T}^n = (b_n/n)\tau^n$ be the associated real tree rescaled so that all edges have length b_n/n (where b_n is the normalizing sequence in (4.1)). Then we have the convergence in distribution $\Psi_{\mathcal{T}^n} \xrightarrow{(d)} \Psi_{\mathcal{T}}$ in $\mathcal{M}(\mathbb{T}\times\mathbb{R}_+)$, where $\mathcal{T}$ is the stable Lévy tree with branching mechanism $\psi(\lambda) = \kappa\lambda^\gamma$. In particular, we have $\Psi_{\mathcal{T}^n}^{\mathfrak{m}\mathfrak{h}} \xrightarrow{(d)} \Psi_{\mathcal{T}^n}^{\mathfrak{m}\mathfrak{h}}$ in $\mathcal{M}(\mathbb{R}_+^2)$.*

The convergence in distribution obtained in Corollary 4.7 is unsatisfactory to study the asymptotics of additive functionals of large BGW trees as it involves the real tree $\mathcal{T}^n$ instead of the (discrete) BGW tree τ^n. To remedy this, we shall introduce a discrete version of the measure Ψ_T when T is associated with a discrete tree. Let $\mathbf{t}$ be a discrete tree and $a > 0$. Recall that $a\mathbf{t}$ denotes the real tree associated to $\mathbf{t}$ where the branches have length a, and that for $v \in \mathbf{t}$, av denotes the corresponding vertex in $a\mathbf{t}$, see Section 2.3 for the definitions. We define two nonnegative measures $\mathcal{A}^\circ_{\mathbf{t},a}$ and $\mathcal{A}_{\mathbf{t},a}$ on $\mathbb{T}\times\mathbb{R}_+$ by, for every $f \in \mathcal{B}_+(\mathbb{T}\times\mathbb{R}_+)$,

$$\boxed{\mathcal{A}^\circ_{\mathbf{t},a}(f) = \frac{a}{|\mathbf{t}|}\sum_{w\in\mathbf{t}^\circ} |\mathbf{t}_w| f\left(a\mathbf{t}_w, aH(w)\right)} \quad \text{and} \quad \boxed{\mathcal{A}_{\mathbf{t},a}(f) = \frac{a}{|\mathbf{t}|}\sum_{w\in\mathbf{t}} |\mathbf{t}_w| f\left(a\mathbf{t}_w, aH(w)\right)}, \tag{4.16}$$

where $a\mathbf{t}_w$ is the subtree of $a\mathbf{t}$ above aw. Note that the sum is over all internal vertices of $\mathbf{t}$ for $\mathcal{A}^\circ_{\mathbf{t},a}$, while for $\mathcal{A}_{\mathbf{t},a}$ the sum extends over all vertices including the leaves. In other words, the measure $\mathcal{A}^\circ_{\mathbf{t},a}$ ignores the subtrees rooted at a leaf of $\mathbf{t}$ (which are trivial trees consisting only of a root equipped with a scaled Dirac measure). Let us take a moment to explain why we introduce the measure $\mathcal{A}^\circ_{\mathbf{t},a}$. While $\mathcal{A}_{\mathbf{t},a}$ seems more natural, the measure $\mathcal{A}^\circ_{\mathbf{t},a}$ has the advantage of putting no mass on the set

$$\mathbb{T}_0 \times \mathbb{R}_+ = \{T \in \mathbb{T}\colon\ \mathfrak{m}(T) = 0 \text{ or } \mathfrak{h}(T) = 0\} \times \mathbb{R}_+.$$

This will be useful as we are interested in sums of the form (4.16) where the function f may blow up on $\mathbb{T}_0\times\mathbb{R}_+$. We now give estimates for the distances between the three measures $\mathcal{A}^\circ_{\mathbf{t},a}$, $\mathcal{A}_{\mathbf{t},a}$ and $\Psi_{a\mathbf{t}}$, on $\mathbb{T}\times\mathbb{R}_+$, which are associated with the discrete tree $\mathbf{t}$ and $a > 0$.

Lemma 4.8. *Let $\mathbf{t}$ be a discrete tree and let $a > 0$. We have*

$$d_{\mathrm{BL}}\left(\Psi_{a\mathbf{t}}, \mathcal{A}_{\mathbf{t},a}\right) \le a\left(\frac{3}{4}\mathcal{A}_{\mathbf{t},a}(1) + 1\right), \tag{4.17}$$

$$d_{\mathrm{TV}}(\mathcal{A}_{\mathbf{t},a}, \mathcal{A}^\circ_{\mathbf{t},a}) \le \frac{1}{2}a. \tag{4.18}$$

Proof. Let $f \in \mathcal{C}_b(\mathbb{T} \times \mathbb{R}_+)$ be Lipschitz. Recall that $T = a\mathbf{t}$ is the real tree associated with $\mathbf{t}$, rescaled so that all edges have length a and equipped with the uniform probability measure on the set of vertices whose height is an integer multiple of a. Recall also that for $v \in \mathbf{t}$, av denotes the corresponding vertex in $T = a\mathbf{t}$. In particular, $H(av) = aH(v)$, where $H(av)$ is the height of av in the real tree $a\mathbf{t}$ and $H(v)$ is the height of v in the discrete tree $\mathbf{t}$. Thus, we have

$$\begin{aligned}\Psi_T(f) = \frac{1}{|\mathbf{t}|}\sum_{v\in\mathbf{t}}\int_0^{H(av)} f(T_{r,av}, r)\,\mathrm{d}r &= \frac{1}{|\mathbf{t}|}\sum_{v\in\mathbf{t}}\int_0^{aH(v)} f(T_{r,av}, r)\,\mathrm{d}r\\ &= \frac{a}{|\mathbf{t}|}\sum_{v\in\mathbf{t}}\sum_{k=1}^{H(v)}\int_{k-1}^{k} f\left(T_{ar,av}, ar\right)\,\mathrm{d}r.\end{aligned}$$

On the other hand, note that for every $1 \leq k \leq H(v)$, we have $T_{ak,av} = T_{aw}$ where $w \in \mathbf{t}$ is the unique ancestor of v with height k. Thus, we have

$$\sum_{v\in\mathbf{t}}\sum_{k=1}^{H(v)} f\left(T_{ak,av}, ak\right) = \sum_{v\in\mathbf{t}}\sum_{\substack{w\preccurlyeq v\\ w\neq\emptyset}} f\left(T_{aw}, aH(w)\right) = \sum_{w\neq\emptyset}|\mathbf{t}_w| f\left(T_{aw}, aH(w)\right) = \frac{|\mathbf{t}|}{a}\mathcal{A}_{\mathbf{t},a}(f) - |\mathbf{t}| f\left(T, 0\right).$$

Therefore, we deduce that

$$\begin{aligned}\left|\Psi_T(f) - \mathcal{A}_{\mathbf{t},a}(f)\right| &\leq \frac{a}{|\mathbf{t}|}\sum_{v\in\mathbf{t}}\sum_{k=1}^{H(v)}\int_{k-1}^{k}\left|f\left(T_{ar,av}, ar\right) - f\left(T_{ak,av}, ak\right)\right|\,\mathrm{d}r + a\left\|f\right\|_\infty\\ &\leq \frac{a}{|\mathbf{t}|}\sum_{v\in\mathbf{t}}\sum_{k=1}^{H(v)}\int_{k-1}^{k}\left\|f\right\|_{\mathrm{L}}\Big(d_{\mathrm{GHP}}\left(T_{ar,av}, T_{ak,av}\right) + a(k-r)\Big)\,\mathrm{d}r + a\left\|f\right\|_\infty. \quad (4.19)\end{aligned}$$

Since for $k-1 < r \leq k$, the tree $T_{ar,av}$ is obtained by grafting $T_{ak,av}$ on top of a branch of height $a(k-r)$ and no mass, it is straightforward to check that $d_{\mathrm{GHP}}\left(T_{ar,av}, T_{ak,av}\right) \leq a(k-r)/2$. It follows that

$$\left|\Psi_T(f) - \mathcal{A}_{\mathbf{t},a}(f)\right| \leq \frac{a}{|\mathbf{t}|}\sum_{v\in\mathbf{t}}\sum_{k=1}^{H(v)}\frac{3a}{4}\left\|f\right\|_{\mathrm{L}} + a\left\|f\right\|_\infty \leq \frac{3a}{4}\left\|f\right\|_{\mathrm{L}}\mathcal{A}_{\mathbf{t},a}(1) + a\left\|f\right\|_\infty.$$

By definition of the distance d_{BL}, we deduce that

$$d_{\mathrm{BL}}\left(\Psi_T, \mathcal{A}_{\mathbf{t},a}\right) \leq a\left(\frac{3}{4}\mathcal{A}_{\mathbf{t},a}(1) + 1\right).$$

Next, let $f \in \mathcal{B}_b(\mathbb{T} \times \mathbb{R}_+)$. We have

$$\left|\mathcal{A}_{\mathbf{t},a}(f) - \mathcal{A}^\circ_{\mathbf{t},a}(f)\right| = \frac{a}{|\mathbf{t}|}\left|\sum_{w\in\mathrm{Lf}(\mathbf{t})}|\mathbf{t}_w| f\left(T_{aw}, aH(w)\right)\right| \leq \frac{a}{|\mathbf{t}|}\left|\mathrm{Lf}(\mathbf{t})\right|\left\|f\right\|_\infty \leq a\left\|f\right\|_\infty.$$

Taking the supremum over all $f \in \mathcal{B}_b(\mathbb{T}\times\mathbb{R}_+)$ such that $\left\|f\right\|_\infty \leq 1$ yields $d_{\mathrm{TV}}\left(\mathcal{A}_{\mathbf{t},a}, \mathcal{A}^\circ_{\mathbf{t},a}\right) \leq \frac{1}{2}a$. □

We now restate the convergence of Corollary 4.7 in terms of the discrete trees τ^n. To avoid cumbersome notations, we write

$$\boxed{\mathcal{A}^\circ_n = \mathcal{A}^\circ_{\tau^n, b_n/n}} \quad \text{and} \quad \boxed{\mathcal{A}_n = \mathcal{A}_{\tau^n, b_n/n}}.$$

Recall that for a discrete tree $\mathbf{t}$, $w \in \mathbf{t}$ and $a > 0$, we have that $\mathfrak{h}(a\mathbf{t}_w) = a\mathfrak{h}(\mathbf{t}_w)$ and $\mathfrak{m}(a\mathbf{t}_w) = |\mathbf{t}_w|/|\mathbf{t}|$. We shall also consider the following variant of the measure $\mathcal{A}_n^\circ$ for functions depending only on the mass and height: for every measurable function f belonging to $\mathcal{B}_+([0,1] \times \mathbb{R}_+)$,

$$\boxed{\mathcal{A}_n^{\mathfrak{mh},\circ}(f) = \frac{b_n}{n^2} \sum_{w \in \tau^{n,\circ}} |\tau_w^n| f\left(\frac{|\tau_w^n|}{n}, \frac{b_n}{n}\mathfrak{h}(\tau_w^n)\right)}. \tag{4.20}$$

We have the following upper bound of their total mass.

Lemma 4.9. *We have:*

$$\mathcal{A}_n^\circ(1) \leq \frac{b_n}{n}\mathfrak{h}(\tau^n) \quad \textit{and} \quad \mathcal{A}_n(1) \leq \frac{b_n}{n}\left(\mathfrak{h}(\tau^n) + 1\right). \tag{4.21}$$

Proof. The proof is elementary as

$$\mathcal{A}_n^\circ(1) = \frac{b_n}{n^2} \sum_{w \in \tau^{n,\circ}} |\tau_w^n| = \frac{b_n}{n^2} \sum_{w \in \tau^{n,\circ}} \sum_{w \preccurlyeq v} 1 \leq \frac{b_n}{n^2} \sum_{v \in \tau^n} \mathfrak{h}(\tau^n) \leq \frac{b_n}{n}\mathfrak{h}(\tau^n),$$

$$\mathcal{A}_n(1) = \frac{b_n}{n^2} \sum_{w \in \tau^n} |\tau_w^n| = \mathcal{A}_n^\circ(1) + \frac{b_n}{n^2}|\mathrm{Lf}(\mathbf{t})| \leq \frac{b_n}{n}\left(\mathfrak{h}(\tau^n) + 1\right).$$

□

We have the following convergence of $\mathcal{A}_n^\circ$ as n goes to infinity.

Corollary 4.10. *Assume that ξ satisfies (ξ1) and (ξ2) and let τ^n be a BGW(ξ) tree conditioned to have n vertices. Then for every $f \in \mathcal{C}_b(\mathbb{T} \times \mathbb{R}_+)$, we have the convergence in distribution and of all positive moments*

$$\mathcal{A}_n^\circ(f) = \frac{b_n}{n^2} \sum_{w \in \tau^{n,\circ}} |\tau_w^n| f\left(\frac{b_n}{n}\tau_w^n, \frac{b_n}{n}H(w)\right) \xrightarrow[n\to\infty]{(d)+\text{moments}} \Psi_{\mathcal{T}}(f), \tag{4.22}$$

where $\mathcal{T}$ is the stable Lévy tree with branching mechanism $\psi(\lambda) = \kappa\lambda^\gamma$. In particular, for every $f \in \mathcal{C}_b([0,1] \times \mathbb{R}_+)$, we have

$$\mathcal{A}_n^{\mathfrak{mh},\circ}(f) = \frac{b_n}{n^2} \sum_{w \in \tau^{n,\circ}} |\tau_w^n| f\left(\frac{|\tau_w^n|}{n}, \frac{b_n}{n}\mathfrak{h}(\tau_w^n)\right) \xrightarrow[n\to\infty]{(d)+\text{moments}} \Psi_{\mathcal{T}}^{\mathfrak{mh}}(f). \tag{4.23}$$

Remark 4.11. By (4.18), we have that a.s. and in L^1

$$d_{\mathrm{TV}}\left(\mathcal{A}_n, \mathcal{A}_n^\circ\right) \xrightarrow[n\to\infty]{} 0.$$

In particular, the convergences of Corollary 4.10 still hold if we sum over τ^n instead of $\tau^{n,\circ}$.

Remark 4.12. Another model of random trees is the class of Pólya trees which are random uniform unordered trees. In [36], Panagiotou and Stufler show that the scaling limit of Pólya trees is the Brownian tree and that the sub-exponential tail bounds of Lemma 4.2 hold in this case with $\alpha = \beta = 2$. Let $\Omega \subset \mathbb{N}$ be such that $\Omega \cap \{0,1\} \neq \Omega$ and let T^n denote the uniform random unordered tree with n vertices and vertex outdegree in Ω. Then there exists a finite constant $c_\Omega > 0$ such that $(c_\Omega/\sqrt{n})\mathsf{T}^n$ converges in distribution to the Brownian tree $\mathcal{T}$ with branching mechanism $\psi(\lambda) = 2\lambda^2$. Thus, the result of Corollary 4.10 holds for T^n and the proof is exactly the same as in the BGW case: for every $f \in \mathcal{C}_b(\mathbb{T} \times \mathbb{R}_+)$,

$$\frac{c_\Omega}{n^{3/2}} \sum_{w \in \mathsf{T}^{n,\circ}} |\mathsf{T}^n_w| f\left(\frac{c_\Omega}{\sqrt{n}} \mathsf{T}^n_w, \frac{c_\Omega}{\sqrt{n}} H(w)\right) \xrightarrow[n\to\infty]{(d)+\text{moments}} \Psi_{\mathcal{T}}(f).$$

Proof of Corollary 4.10. Denote by $\mathcal{T}^n = (b_n/n)\tau^n$ the real tree associated with τ^n rescaled so that all edges have length b_n/n and equipped with the uniform probability measure on the set of vertices whose height is an integer multiple of b_n/n. By Lemma 4.8, we have

$$d_{\mathrm{BL}}\left(\Psi_{\mathcal{T}^n}, \mathcal{A}^\circ_n\right) \leq d_{\mathrm{BL}}\left(\Psi_{\mathcal{T}^n}, \mathcal{A}_n\right) + 2 d_{\mathrm{TV}}(\mathcal{A}_n, \mathcal{A}^\circ_n) \leq \frac{b_n}{n}\left(\frac{3}{4}\mathcal{A}_n(1) + 2\right).$$

Thanks to (4.21) and Lemma 4.4, we have that $M = \sup_{n\in\Delta} \mathbb{E}\left[\mathcal{A}_n(1)\right]$ is finite. It follows that

$$\limsup_{n\to\infty} \mathbb{E}\left[d_{\mathrm{BL}}\left(\Psi_{\mathcal{T}^n}, \mathcal{A}^\circ_n\right)\right] \leq \lim_{n\to\infty} \frac{b_n}{n}\left(\frac{3M}{4} + 2\right) = 0.$$

Thus, using that $\Psi_{\mathcal{T}^n} \xrightarrow{(d)} \Psi_{\mathcal{T}}$ in $\mathcal{M}(\mathbb{T} \times \mathbb{R}_+)$ by Corollary 4.7, Slutsky's lemma yields the convergence in distribution $\mathcal{A}^\circ_n \xrightarrow{(d)} \Psi_{\mathcal{T}}$ in $\mathcal{M}(\mathbb{T} \times \mathbb{R}_+)$ which proves (4.22).

Let $f \in \mathcal{C}_b(\mathbb{T} \times \mathbb{R}_+)$. Using Skorokhod's representation theorem, we may assume that the convergence (4.22) holds almost surely. To prove the convergence of positive moments, it suffices to show that the family $(\mathcal{A}^\circ_n(f),\, n \in \Delta)$ is bounded in L^p for every $p \in [1, \infty)$. This is the case as by (4.21), we have $\mathcal{A}^\circ_n(f) \leq \|f\|_\infty \mathcal{A}^\circ_n(1) \leq \|f\|_\infty \frac{b_n}{n}\mathfrak{h}(\tau^n)$, and the family $(\frac{b_n}{n}\mathfrak{h}(\tau^n),\, n \in \Delta)$ is bounded in L^p for every $p \in [1, \infty)$ by Lemma 4.4. This completes the proof. □

The Gromov-Hausdorff-Prokhorov convergence (4.15) allowed us to derive an invariance principle (4.22) for a certain class of additive functionals on BGW trees, namely those associated with real-valued continuous bounded functions f defined on $\mathbb{T} \times \mathbb{R}_+$. In the sequel, we will be looking at a similar invariance principle when f blows up on $\mathbb{T}_0 \times \mathbb{R}_+$. It is not surprising that the Gromov-Hausdorff-Prokhorov convergence alone does not allow us to say anything about the convergence of $\Psi_{\mathcal{T}^n}(f)$ in this case as the next remark illustrates.

Remark 4.13. Let τ^n be a Catalan tree with n vertices, where $n \in \Delta = 2\mathbb{N} + 1$. In other words, τ^n is uniformly distributed among the set of full binary ordered trees with n vertices, which corresponds to a BGW(ξ) tree with $\mathbb{P}(\xi = 0) = \mathbb{P}(\xi = 2) = 1/2$ conditioned to have size n. Notice that ξ has finite variance $\sigma^2_\xi = 1$. Take $b_n = \sqrt{n}/2$ so that by (4.15), $\mathcal{T}^n = (1/2\sqrt{n})\tau^n$ converges in distribution in $\mathbb{T}$ to the Brownian continuum random tree $\mathcal{T}$ with branching mechanism $\psi(\lambda) = 2\lambda^2$. In fact, it is well known, see *e.g.* [37, Theorem 7.9], that there is a representation of $\mathcal{T}^n$ such that the almost sure convergence holds. Denote by $\mathcal{T}^n_\varepsilon$ the real tree obtained from $\mathcal{T}^n$ by stretching the leaves by a distance of $\varepsilon \geq 0$ and equip it with the uniform probability measure on the set of branching points and leaves. Fix $0 < \alpha < 1/2$ and set $\varepsilon_n = n^{-\alpha}$. It is clear from this construction that $\mathcal{T}^n_{\varepsilon_n}$ is a $\mathbb{T}$-valued random variable and that a.s.

$$d_{\mathrm{GHP}}\left(\mathcal{T}^n_{\varepsilon_n}, \mathcal{T}^n\right) \leq \varepsilon_n.$$

So it follows that $\mathcal{T}^n_{\varepsilon_n}$ converges to $\mathcal{T}$ a.s. in the sense of the Gromov-Hausdorff-Prokhorov distance. We consider $f(T, r) = \mathfrak{m}(T)^{-\alpha}$ and if $\nu \in \mathcal{M}(\mathbb{T} \times \mathbb{R}_+)$ we write $\nu(x^{-\alpha})$ for $\nu(f)$. According to

[14, Theorem 3.1], we have the following a.s. convergence $\mathcal{A}_n(x^{-\alpha}) \xrightarrow[n\to\infty]{} \Psi_{\mathcal{T}}(x^{-\alpha})$. In conjunction with the identity $\Psi_{\mathcal{T}^n}(x^{-\alpha}) = \mathcal{A}_n(x^{-\alpha}) - 1/(2\sqrt{n})$ this proves the a.s. convergence

$$\Psi_{\mathcal{T}^n}(x^{-\alpha}) \xrightarrow[n\to\infty]{} \Psi_{\mathcal{T}}(x^{-\alpha}).$$

On the other hand, we have

$$\Psi_{\mathcal{T}^n_{\varepsilon_n}}(x^{-\alpha}) - \Psi_{\mathcal{T}^n}(x^{-\alpha}) = \frac{1}{|\tau^n|} \sum_{w\in \mathrm{Lf}(\tau^n)} \int_{(2\sqrt{n})^{-1}H(w)}^{(2\sqrt{n})^{-1}H(w)+\varepsilon_n} \left(\frac{|\tau^n_w|}{|\tau^n|}\right)^{-\alpha} \mathrm{d}r = \frac{n+1}{2} n^{\alpha-1}\varepsilon_n$$

since $|\tau^n| = n$ and $|\,\mathrm{Lf}(\tau^n)| = (n+1)/2$. Thus, we get

$$\Psi_{\mathcal{T}^n_{\varepsilon_n}}(x^{-\alpha}) - \Psi_{\mathcal{T}^n}(x^{-\alpha}) \xrightarrow[n\to\infty]{} \frac{1}{2}.$$

In conclusion, even though we have the a.s. convergence $\mathcal{T}^n_{\varepsilon_n}$ towards $\mathcal{T}$ in $\mathbb{T}$, $\Psi_{\mathcal{T}^n_{\varepsilon_n}}(x^{-\alpha})$ does not converge to $\Psi_{\mathcal{T}}(x^{-\alpha})$ for $\alpha \in (0, 1/2)$. This proves that the continuity of $\Psi_T(f)$ in T when f blows up on $\mathbb{T}_0$, which has been observed in [14], is indeed specific to BGW trees.

5. Technical lemmas

In this section, we gather some technical results that will be used later. The next lemma, which gives sufficient conditions for boundedness in L^1 of functionals of the mass and height on BGW trees, will be a key ingredient in proving our convergence results. Recall that τ is a BGW(ξ) tree and τ^n is a BGW(ξ) conditioned to have n vertices. Recall from (4.20) the definition of the measure $\mathcal{A}_n^{\mathfrak{mh},\circ}$ and notice that $\mathcal{A}_n^{\mathfrak{mh},\circ}([0,1]\times\mathbb{R}_+ \setminus (0,1]\times\mathbb{R}_+^*) = 0$. For this reason, we also see $\mathcal{A}_n^{\mathfrak{mh},\circ}$ as a measure on $(0,1]\times\mathbb{R}_+^*$. By convention, we write $\mathcal{A}_n^{\mathfrak{mh},\circ}(g(x)h(u))$ for $\mathcal{A}_n^{\mathfrak{mh},\circ}(f)$ where $f(x,u) = g(x)h(u)$, and we see g as a function of the mass and h as a function of the height.

Lemma 5.1. *Assume that ξ satisfies (ξ1) and (ξ2)′. Suppose that $f \in \mathcal{B}_+((0,1]\times\mathbb{R}_+^*)$ satisfies one of the following assumptions:*

(i) f is of the form $f(x,u) = g(x)u^\beta$ or $f(x,u) = x^\alpha h(u)$ where $\alpha, \beta \in \mathbb{R}$ and g, h are nonincreasing and

$$\int_0 f(x^{\gamma/(\gamma-1)}, x)\,\mathrm{d}x < \infty. \tag{5.1}$$

(ii) $f(x,u) = g(x)\mathrm{e}^{u^\eta}\mathbf{1}_{[1,\infty)}(u)$ where $\eta \in (0,\gamma)$ and $g \in \mathcal{B}_+((0,1])$ is nonincreasing and satisfies $\int_0 g(x)\mathrm{e}^{-x^{-r_0}}\,\mathrm{d}x < \infty$ for some $r_0 \in (0, \gamma-1)$.

Then, we have

$$\sup_{n\in\Delta} \mathbb{E}\left[\mathcal{A}_n^{\mathfrak{mh},\circ}(f)\right] < \infty.$$

Proof of Lemma 5.1. Here c, C and M denote positive finite constants that may vary from expression to expression (but are independent of n and x). Let $n \in \Delta$ so that $\mathbb{P}(S_n = n-1) > 0$. Observe that $w \in \tau^{n,\circ}$ if and only if $|\tau^n_w| > 1$ and that the root $\emptyset$ is the only vertex in τ^n such that $|\tau^n_w| = n$. Thus, for every $f \in \mathcal{B}_+([0,1]\times\mathbb{R}_+)$, we have the decomposition

$$\mathbb{E}\left[\mathcal{A}_n^{\mathfrak{mh},\circ}(f)\right] = \frac{b_n}{n^2}\,\mathbb{E}\left[\sum_{w\in\tau^{n,\circ}} |\tau^n_w| f\left(\frac{|\tau^n_w|}{n}, \frac{b_n}{n}\mathfrak{h}(\tau^n_w)\right)\right]$$

$$= \frac{b_n}{n^2} \mathbb{E}\left[\sum_{w\in\tau^n} \mathbf{1}_{\{1<|\tau^n_w|<n\}} |\tau^n_w| f\left(\frac{|\tau^n_w|}{n}, \frac{b_n}{n}\mathfrak{h}(\tau^n_w)\right)\right] + \frac{b_n}{n}\mathbb{E}\left[f\left(1, \frac{b_n}{n}\mathfrak{h}(\tau^n)\right)\right].$$

By [29, Lemma 5.1], we have

$$\frac{b_n}{n^2} \mathbb{E}\left[\sum_{w\in\tau^n} \mathbf{1}_{\{1<|\tau^n_w|<n\}} |\tau^n_w| f\left(\frac{|\tau^n_w|}{n}, \frac{b_n}{n}\mathfrak{h}(\tau^n_w)\right)\right]$$
$$= \frac{b_n}{n}\sum_{k=1}^n \frac{\mathbb{P}(S_k = k-1)\,\mathbb{P}(S_{n-k}=n-k)}{\mathbb{P}(S_n = n-1)} \mathbb{E}\left[f\left(\frac{k}{n}, \frac{b_n}{n}\mathfrak{h}(\tau^k)\right)\right]\mathbf{1}_{\{1<k<n\}}, \quad (5.2)$$

where by convention the summand is zero for $k \notin \Delta$. Using Lemma 4.1 and (4.2), we get for every $n \in \Delta$ and every $1 < k < n$

$$b_n \frac{\mathbb{P}(S_k = k-1)\,\mathbb{P}(S_{n-k}=n-k)}{\mathbb{P}(S_n = n-1)} \leq C\frac{b_n^2}{b_k b_{n-k}} \leq C\left(\frac{n^2}{k(n-k)}\right)^{1/\gamma}.$$

We deduce that

$$\frac{b_n}{n^2}\mathbb{E}\left[\sum_{w\in\tau^{n,\circ}} |\tau^n_w| f\left(\frac{|\tau^n_w|}{n}, \frac{b_n}{n}\mathfrak{h}(\tau^n_w)\right)\right] \leq \frac{C}{n}\sum_{k=1}^n g_n(k) + \frac{b_n}{n}\mathbb{E}\left[f\left(1, \frac{b_n}{n}\mathfrak{h}(\tau^n)\right)\right]$$
$$= C\int_0^1 g_n(\lceil nx\rceil)\,\mathrm{d}x + \frac{b_n}{n}\mathbb{E}\left[f\left(1, \frac{b_n}{n}\mathfrak{h}(\tau^n)\right)\right], \quad (5.3)$$

where we set

$$g_n(k) = \left(\frac{n^2}{k(n-k)}\right)^{1/\gamma}\mathbb{E}\left[f\left(\frac{k}{n}, \frac{b_n}{n}\mathfrak{h}(\tau^k)\right)\right]\mathbf{1}_{\{1<k<n\}} \quad \text{for all } k\in\Delta, \quad (5.4)$$

and $g_n(k) = 0$ for $k \notin \Delta$. We will constantly make use of the following inequality

$$c\left(\frac{k}{n}\right)^{1-1/\gamma} \leq \frac{b_n}{n}\frac{k}{b_k} \leq C\left(\frac{k}{n}\right)^{1-1/\gamma} \quad \text{for all } 1\leq k\leq n, \quad (5.5)$$

which follows easily from (4.2).

First case. Assume (i). First, we consider the case $f(x,u) = g(x)u^\beta$. Since $b_n/n \to 0$, we deduce from Lemma 4.4 that

$$\lim_{n\to\infty}\frac{b_n}{n}\mathbb{E}\left[f\left(1, \frac{b_n}{n}\mathfrak{h}(\tau^n)\right)\right] = g(1)\lim_{n\to\infty}\frac{b_n}{n}\mathbb{E}\left[\left(\frac{b_n}{n}\mathfrak{h}(\tau^n)\right)^\beta\right] = 0. \quad (5.6)$$

For every $1/n < x \leq (n-1)/n$, it holds that $x \leq \lceil nx\rceil/n \leq 2x$ and $n - \lceil nx\rceil \geq n(1-x)/2$. Thus, for every $x \in (0,1)$, using Lemma 4.4 for the last inequality, we have

$$g_n(\lceil nx\rceil) \leq Mx^{-1/\gamma}(1-x)^{-1/\gamma} g\left(\frac{\lceil nx\rceil}{n}\right)\mathbb{E}\left[\left(\frac{b_n}{n}\mathfrak{h}(\tau^{\lceil nx\rceil})\right)^\beta\right]\mathbf{1}_{\{1<nx\leq n-1\}}$$
$$\leq Mx^{-1/\gamma}(1-x)^{-1/\gamma}g(x)\left(\frac{b_n}{n}\frac{\lceil nx\rceil}{b_{\lceil nx\rceil}}\right)^\beta \sup_{k\in\Delta}\mathbb{E}\left[\left(\frac{b_k}{k}\mathfrak{h}(\tau^k)\right)^\beta\right]\mathbf{1}_{\{1<nx\leq n-1\}}$$
$$\leq Mx^{(\beta+1)(1-1/\gamma)-1}(1-x)^{-1/\gamma}g(x).$$

It follows that

$$\int_0^1 g_n(\lceil nx\rceil)\,\mathrm{d}x \leq M\int_0^1 g(x)x^{(\beta+1)(1-1/\gamma)-1}(1-x)^{-1/\gamma}\,\mathrm{d}x, \quad (5.7)$$

where the right-hand side is finite by (5.1) as $\gamma > 1$. Combining (5.6) and (5.7), it follows from (5.3) that

$$\sup_{n\in\Delta} \mathbb{E}\left[\mathcal{A}_n^{\mathfrak{m}\mathfrak{h},\circ}(f)\right] = \sup_{n\in\Delta} \frac{b_n}{n^2}\,\mathbb{E}\left[\sum_{w\in\tau^{n,\circ}} |\tau_w^n| f\left(\frac{|\tau_w^n|}{n}, \frac{b_n}{n}\mathfrak{h}(\tau_w^n)\right)\right] < \infty.$$

Next, we consider the case $f(x,u) = x^\alpha h(u)$. By Lemma 4.2 and (i) from Remark 4.3, we have, for every $k \in \Delta$,

$$\mathbb{P}\left(\frac{b_k}{k}\mathfrak{h}(\tau^k) \le y\right) \le 1 \wedge \left(C_0 \exp\left(-c_0 y^{-\gamma/(\gamma-1)}\right)\right). \tag{5.8}$$

Denoting by Y a random variable whose cdf is given by the right-hand side and using (5.5), we get, for every $2 \le k \le n$,

$$\frac{b_n}{n}\mathfrak{h}(\tau^k) \ge_{\mathrm{st}} \frac{b_n}{n}\frac{k}{b_k} Y \ge c\left(\frac{k}{n}\right)^{1-1/\gamma} Y, \tag{5.9}$$

where $\ge_{\mathrm{st}}$ denotes the usual stochastic order. In particular, since Y has density

$$y \mapsto C y^{-(2\gamma-1)/(\gamma-1)} \exp\left(-c_0 y^{-\gamma/(\gamma-1)}\right) \mathbf{1}_{[0,a]}(y)$$

for some $a > 0$, the first inequality in (5.9) applied with $k = n$ gives, for every $n \in \Delta$,

$$\mathbb{E}\left[h\left(\frac{b_n}{n}\mathfrak{h}(\tau^n)\right)\right] \le \mathbb{E}\left[h(Y)\right] \le C\int_0^\infty h(y)\mathrm{e}^{-c_0 y^{-\gamma/(\gamma-1)}} \frac{\mathrm{d}y}{y^{(2\gamma-1)/(\gamma-1)}}. \tag{5.10}$$

Note that the last integral is finite: indeed, since h is nonincreasing, we have

$$\int_1^\infty h(y)\mathrm{e}^{-c_0 y^{-\gamma/(\gamma-1)}} \frac{\mathrm{d}y}{y^{(2\gamma-1)/(\gamma-1)}} \le h(1)\int_1^\infty \frac{\mathrm{d}y}{y^{(2\gamma-1)/(\gamma-1)}} < \infty,$$

and by (5.1)

$$\int_0^1 h(y)\mathrm{e}^{-c_0 y^{-\gamma/(\gamma-1)}} \frac{\mathrm{d}y}{y^{(2\gamma-1)/(\gamma-1)}} \le \sup_{0<y\le 1} \frac{\mathrm{e}^{-c_0 y^{-\gamma/(\gamma-1)}}}{y^{1+(\alpha+1)\gamma/(\gamma-1)}} \int_0^1 h(y) y^{\alpha\gamma/(\gamma-1)}\,\mathrm{d}y < \infty. \tag{5.11}$$

Then, applying (5.9) with $k = \lceil nx \rceil$ and using the fact that h is nonincreasing, we get for every $x \in (0,1)$

$$\begin{aligned}
g_n(\lceil nx\rceil) &\le M x^{-1/\gamma}(1-x)^{-1/\gamma}\left(\frac{\lceil nx\rceil}{n}\right)^\alpha \mathbb{E}\left[h\left(\frac{b_n}{n}\mathfrak{h}(\tau^{\lceil nx\rceil})\right)\right]\mathbf{1}_{\{1<nx\le n-1\}}\\
&\le M x^{\alpha-1/\gamma}(1-x)^{-1/\gamma}\,\mathbb{E}\left[h\left(c x^{1-1/\gamma} Y\right)\right]\\
&\le M x^{\alpha-1/\gamma}(1-x)^{-1/\gamma}\int_0^a h\left(c x^{1-1/\gamma} y\right)\mathrm{e}^{-c_0 y^{-\gamma/(\gamma-1)}} \frac{\mathrm{d}y}{y^{(2\gamma-1)/(\gamma-1)}}\\
&\le M x^{1+\alpha-1/\gamma}(1-x)^{-1/\gamma}\int_0^{acx^{1-1/\gamma}} h(u)\mathrm{e}^{-rxu^{-\gamma/(\gamma-1)}} \frac{\mathrm{d}u}{u^{(2\gamma-1)/(\gamma-1)}},
\end{aligned}$$

for some positive constant $r > 0$, where in the last inequality we made the change of variable $u = cx^{1-1/\gamma}y$. Therefore we have

$$\int_0^1 g_n\left(\lceil nx\rceil\right)\mathrm{d}x \le M\int_0^1 x^{1+\alpha-1/\gamma}(1-x)^{-1/\gamma}\,\mathrm{d}x \int_0^{acx^{1-1/\gamma}} h(u)\mathrm{e}^{-rxu^{-\gamma/(\gamma-1)}} \frac{\mathrm{d}u}{u^{(2\gamma-1)/(\gamma-1)}}. \tag{5.12}$$

It remains to check that the last integral is finite. But, arguing as in (5.11) with r instead of c_0, we have

$$\int_{1/2}^1 x^{1+\alpha-1/\gamma}(1-x)^{-1/\gamma}\,\mathrm{d}x\int_0^{acx^{1-1/\gamma}} h(u)\mathrm{e}^{-rxu^{-\gamma/(\gamma-1)}}\,\frac{\mathrm{d}u}{u^{(2\gamma-1)/(\gamma-1)}}$$
$$\leq M\int_{1/2}^1 (1-x)^{-1/\gamma}\,\mathrm{d}x\int_0^{ac} h(u)\mathrm{e}^{-ru^{-\gamma/(\gamma-1)}/2}\,\frac{\mathrm{d}u}{u^{(2\gamma-1)/(\gamma-1)}}<\infty.$$

Let $\delta=\gamma/(\gamma-1)$. Making the change of variable $y=xu^{-\delta}$ with u fixed, we have, thanks to (5.1),

$$\begin{aligned}\int_0^{1/2} x^{1+\alpha-1/\gamma}(1-x)^{-1/\gamma}\,\mathrm{d}x\int_0^{acx^{1-1/\gamma}} h(u)\mathrm{e}^{-rxu^{-\delta}}\,\frac{\mathrm{d}u}{u^{1+\delta}}&\\ \leq \int_{(ac)^{-\delta}}^\infty y^{1+\alpha-1/\gamma}\mathrm{e}^{-ry}\,\mathrm{d}y\int_0^\infty h(u)u^{\alpha\delta}\mathbf{1}_{\{yu^\delta\leq 1/2\}}\,\mathrm{d}u&\\ \leq \int_{(ac)^{-\delta}}^\infty y^{1+\alpha-1/\gamma}\mathrm{e}^{-ry}\,\mathrm{d}y\int_0^{ac} h(u)u^{\alpha\delta}\,\mathrm{d}u<\infty.&\end{aligned}$$

The right-hand side of (5.10) and (5.12) being finite and $(b_n/n,\, n\geq 1)$ being bounded, we deduce from (5.3) that

$$\sup_{n\in\Delta}\mathbb{E}\left[\mathcal{A}_n^{\mathfrak{m}\mathfrak{h},\circ}(f)\right]=\sup_{n\in\Delta}\frac{b_n}{n^2}\,\mathbb{E}\left[\sum_{w\in\tau^{n,\circ}}|\tau_w^n|f\left(\frac{|\tau_w^n|}{n},\frac{b_n}{n}\mathfrak{h}(\tau_w^n)\right)\right]<\infty.$$

Second case. Assume (ii). Fix $\eta\in(0,\gamma)$ and set $h(u)=\mathrm{e}^{u^\eta}\mathbf{1}_{\{u\geq 1\}}$. Choose $\beta\in(\eta,\gamma)$ such that $\beta(1-1/\gamma)>r_0$. By (4.9) and (5.5), we have, for every $k\in\Delta$ such that $2\leq k\leq n$,

$$\frac{b_n}{n}\mathfrak{h}(\tau^k)\leq_{\mathrm{st}}\frac{b_n}{n}\frac{k}{b_k}Z\leq C\left(\frac{k}{n}\right)^{1-1/\gamma}Z, \tag{5.13}$$

where Z has density $z\mapsto Mz^{\beta-1}\mathrm{e}^{-c_0z^\beta}\mathbf{1}_{[a,\infty)}(z)$ for some $a>0$. So, we get for $x\in(0,1)$

$$\begin{aligned}g_n(\lceil nx\rceil)&\leq Mx^{-1/\gamma}(1-x)^{-1/\gamma}g\left(\frac{\lceil nx\rceil}{n}\right)\mathbb{E}\left[h\left(\frac{b_n}{n}\mathfrak{h}(\tau^{\lceil nx\rceil})\right)\right]\\ &\leq Mx^{-1/\gamma}(1-x)^{-1/\gamma}g\left(x\right)\mathbb{E}\left[h\left(Cx^{1-1/\gamma}Z\right)\right]\\ &\leq Mx^{-1/\gamma}(1-x)^{-1/\gamma}g\left(x\right)\int_a^\infty h\left(Cx^{1-1/\gamma}z\right)z^{\beta-1}\mathrm{e}^{-c_0z^\beta}\,\mathrm{d}z\\ &\leq Mx^{-1/\gamma}(1-x)^{-1/\gamma}g\left(x\right)\int_a^\infty z^{\beta-1}\mathrm{e}^{c_1z^\eta-c_0z^\beta}\mathbf{1}_{\{Cx^{1-1/\gamma}z\geq 1\}}\,\mathrm{d}z,\end{aligned}$$

where we used (5.5) for the first and second inequalities, the monotonicity of g and h for the second and the fact that $\left(Cx^{1-1/\gamma}z\right)^\eta\leq c_1z^\eta$ for some finite constant $c_1>0$ for the last. Notice that if $r<c_0$, then the function $z\mapsto\mathrm{e}^{c_1z^\eta-(c_0-r)z^\beta}$ is bounded on $\mathbb{R}_+$ as $\beta>\eta$. It follows that

$$\begin{aligned}\int_0^1 g_n(\lceil nx\rceil)\,\mathrm{d}x&\leq M\int_0^1 x^{-1/\gamma}(1-x)^{-1/\gamma}g(x)\,\mathrm{d}x\int_0^\infty z^{\beta-1}\mathrm{e}^{-rz^\beta}\mathbf{1}_{\{Cx^{1-1/\gamma}z\geq 1\}}\,\mathrm{d}z\\ &\leq M\int_0^1 x^{-1/\gamma}(1-x)^{-1/\gamma}\mathrm{e}^{-rC^{-\beta}x^{-\beta(1-1/\gamma)}}g(x)\,\mathrm{d}x\\ &\leq M\int_0^1(1-x)^{-1/\gamma}\mathrm{e}^{-x^{-r_0}}g(x)\,\mathrm{d}x<\infty,\end{aligned} \tag{5.14}$$

where in the last inequality we used that the function $x \mapsto x^{-1/\gamma} \mathrm{e}^{x^{-r_0} - rC^{-\beta} x^{-\beta(1-1/\gamma)}}$ is bounded on $(0,1]$ as $\beta(1-1/\gamma) > r_0$. On the other hand, we have

$$\frac{b_n}{n}\,\mathbb{E}\left[f\left(1, \frac{b_n}{n}\mathfrak{h}(\tau^n)\right)\right] \le \frac{b_n}{n} g(1)\,\mathbb{E}\left[h(Z)\right] \le M\frac{b_n}{n}\int_1^\infty z^{\beta-1}\mathrm{e}^{cz^\eta - c_0 z^\beta}\,\mathrm{d}z \le M, \tag{5.15}$$

where we used the first inequality from (5.13) with $k=n$ and the fact that h in nondecreasing for the first inequality and that b_n/n converges to 0 as $n\to\infty$ for the last. Combining (5.14) and (5.15), we deduce from (5.3) that

$$\sup_{n\in\Delta}\mathbb{E}\left[\mathcal{A}_n^{\mathfrak{m}\mathfrak{h},\circ}(f)\right] = \sup_{n\in\Delta}\frac{b_n}{n^2}\,\mathbb{E}\left[\sum_{w\in\tau^{n,\circ}} |\tau_w^n| f\left(\frac{|\tau_w^n|}{n}, \frac{b_n}{n}\mathfrak{h}(\tau_w^n)\right)\right] < \infty.$$

$\square$

As a consequence of the following lemma, we get that $(\mathcal{A}_n^{\mathfrak{m}\mathfrak{h},\circ}(x^\alpha u^\beta),\, n\in\Delta)$ is bounded in L^p for some $p>1$.

Lemma 5.2. *Let $\alpha,\beta\in\mathbb{R}$ such that $\gamma\alpha + (\gamma-1)(\beta+1) > 0$. For every $p\ge 1$ such that $p(\gamma\alpha + (\gamma-1)\beta) > 1-\gamma$ and $\delta\in\mathbb{R}$, we have:*

$$\sup_{n\in\Delta}\mathbb{E}\left[\left(\frac{b_n}{n}\mathfrak{h}(\tau^n)\right)^\delta \mathcal{A}_n^{\mathfrak{m}\mathfrak{h},\circ}(x^\alpha u^\beta)^p\right] < \infty. \tag{5.16}$$

Proof. Set $M_n = \frac{b_n}{n}\mathfrak{h}(\tau^n)$ for $n\in\Delta$. Let $p_0, q_0\in(1,\infty)$ such that $1/p_0 + 1/q_0 = 1$. By Hölder's inequality and thanks to (4.21), we have

$$\mathcal{A}_n^{\mathfrak{m}\mathfrak{h},\circ}(x^\alpha u^\beta)^{p_0} \le M_n^{p_0/q_0}\mathcal{A}_n^{\mathfrak{m}\mathfrak{h},\circ}(x^{p_0\alpha}u^{p_0\beta}). \tag{5.17}$$

Assume that $p_0 > p$ satisfies $p_0(\gamma\alpha + (\gamma-1)\beta) > 1-\gamma$. Set $r = p_0/p$ and s such that $1/r + 1/s = 1$. We deduce that

$$\begin{aligned}
\mathbb{E}\left[M_n^\delta \mathcal{A}_n^{\mathfrak{m}\mathfrak{h},\circ}(x^\alpha u^\beta)^p\right] &= \mathbb{E}\left[M_n^{\delta+p/q_0} M_n^{-p/q_0}\mathcal{A}_n^{\mathfrak{m}\mathfrak{h},\circ}(x^\alpha u^\beta)^p\right] \\
&\le \mathbb{E}\left[M_n^{s(\delta+p/q_0)}\right]^{1/s}\mathbb{E}\left[M_n^{-p_0/q_0}\mathcal{A}_n^{\mathfrak{m}\mathfrak{h},\circ}(x^\alpha u^\beta)^{p_0}\right]^{1/r} \\
&\le \mathbb{E}\left[M_n^{s(\delta+p/q_0)}\right]^{1/s}\mathbb{E}\left[\mathcal{A}_n^{\mathfrak{m}\mathfrak{h},\circ}(x^{p_0\alpha}u^{p_0\beta})\right]^{1/r},
\end{aligned}$$

where we used Hölder's inequality for the first inequality and (5.17) for the second. Since $p_0(\gamma\alpha + (\gamma-1)\beta) > 1-\gamma$, the function $f(x,u) = x^{p_0\alpha}u^{p_0\beta}$ satisfies assumption (i) of Lemma 5.1. We deduce that $\sup_{n\in\Delta}\mathbb{E}\left[\mathcal{A}_n^{\mathfrak{m}\mathfrak{h},\circ}(x^{p_0\alpha}u^{p_0\beta})\right] < \infty$. Then use Lemma 4.4 to get (5.16). $\square$

6. Functionals of the mass and height on the stable Lévy tree

In this section, our goal is to study the finiteness and compute the first moment of the random variable $\Psi_{\mathcal{T}}^{\mathfrak{m}\mathfrak{h}}(f)$ where $\mathcal{T}$ is the stable Lévy tree and f is a measurable function. Recall from Section 4.2 that H denotes the ψ-height process under its excursion measure $\mathbf{N}$, σ is the duration of an excursion and $\mathfrak{h}$ is its height. Notice that σ and $\mathfrak{h}$ are the mass and the height of the tree $\mathcal{T}_H$ coded by H. Furthermore, the stable Lévy tree $\mathcal{T}$ (under $\mathbb{P}$) is the real tree $\mathcal{T}_H$ coded by H, see Remark 2.1, under $\mathbf{N}^{(1)}[\bullet] = \mathbf{N}[\bullet \mid \sigma = 1]$.

6.1. **On the fragmentation (on the skeleton) of Lévy trees.** In this section only we consider a general continuous height process H under its excursion measure $\mathbf{N}$ associated with a branching mechanism $\psi(\lambda) = a\lambda + \beta(\lambda^2/2) + \int \pi(\mathrm{d}r)(\mathrm{e}^{-\lambda r} - 1 + \lambda r)$ with $a, \beta \geq 0$, π a σ-finite measure on $(0, \infty)$ such that $\int \pi(\mathrm{d}r)\,(r \wedge r^2) < \infty$ and such that $\int^\infty \mathrm{d}\lambda/\psi(\lambda) < \infty$. We refer to [16, Section 1] for a complete presentation of the subject.

We will present a decomposition of a general Lévy tree using Bismut's decomposition. Define the length and height of the excursion of H above level r that straddles s

$$\sigma_{r,s} = \int_0^\sigma \mathbf{1}_{\{m(s,t)\geq r\}}\,\mathrm{d}t = \mathrm{T}^+_{r,s} - \mathrm{T}^-_{r,s} \quad \text{and} \quad \mathfrak{h}_{r,s} = \sup_{t\in[\mathrm{T}^-_{r,s},\mathrm{T}^+_{r,s}]} H(t) - r, \tag{6.1}$$

where $m(s,t) = \inf_{[s\wedge t, s\vee t]} H$ is the minimum of H between times s, t and $\mathrm{T}^-_{r,s} = \sup\{t < s\colon\ H(t) = r\}$ and $\mathrm{T}^+_{r,s} = \inf\{t > s\colon\ H(t) = r\}$ are the beginning and the end of the excursion of H above level r that straddles time s, see Figure 1. Then, we consider $H^+_{r,s} = (H^+_{r,s}(t), t \geq 0)$ the excursion of H above level r that straddles s defined as:

$$H^+_{r,s}(t) = H\left((t + \mathrm{T}^-_{r,s}) \wedge \mathrm{T}^+_{r,s}\right) - r,$$

and $H^-_{r,s} = (H^-_{r,s}(t), t \geq 0)$ the excursion of H below defined as $H^-_{r,s}(t) = H(t)$ for $t \in [0, \mathrm{T}^-_{r,s}]$ and $H^-_{r,s}(t + \sigma_{r,s})$ for $t > \mathrm{T}^-_{r,s}$. Notice that the duration and height of the excursion $H^+_{r,s}$ are given by $\sigma^+_{r,s} = \sigma_{r,s}$ and $H_{r,s}$; that the duration of the excursion $H^-_{r,s}$ is given by $\sigma^-_{r,s} = \sigma - \sigma_{r,s}$; and that

$$\sigma = \sigma^+_{r,s} + \sigma^-_{r,s}. \tag{6.2}$$

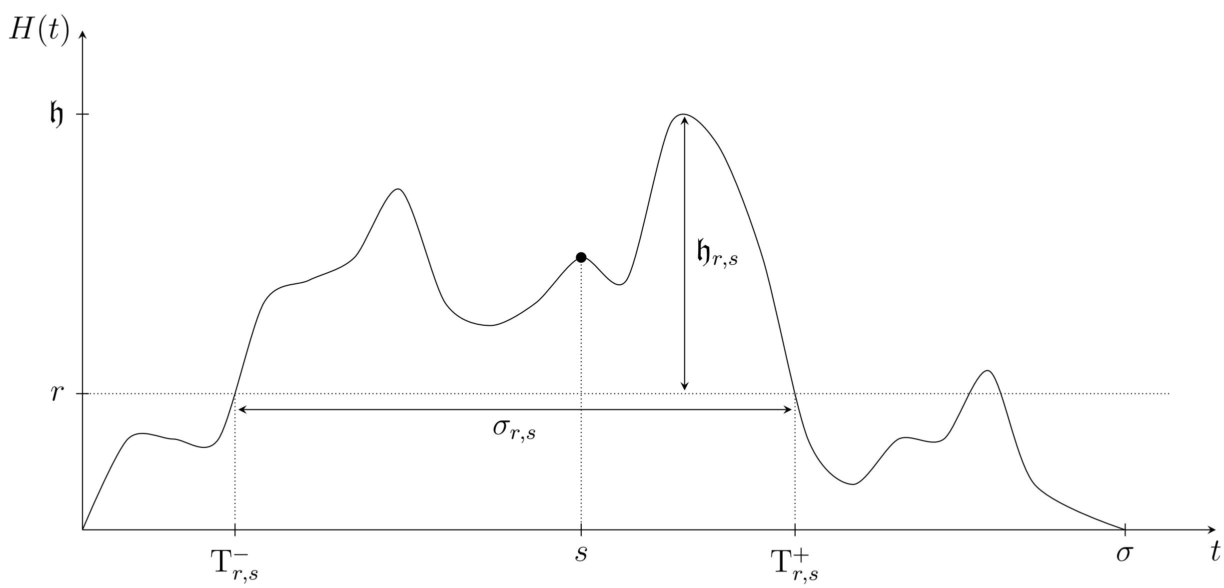

FIGURE 1. The duration $\sigma_{r,s}$ and the height $\mathfrak{h}_{r,s}$ of the excursion of H above level r that straddles time s.

Recall notations from Remark 2.1. For $s \in [0, \sigma]$ and $r \in [0, H(s)]$, the function $H^+_{r,s}$ codes the subtree $\mathcal{T}_{r,\,s} := (\mathcal{T}_H)_{r,\,p(s)}$ and $H^-_{r,s}$ codes the subtree $\mathcal{T}^-_{r,\,s} := (\mathcal{T}_H \setminus \mathcal{T}_{r,\,s}) \cup \{x_{r,s}\}$, where $x_{r,s}$ is the ancestor of $p(s)$, the image of s on $\mathcal{T}_H$, at distance r from the root of $\mathcal{T}_H$. The next lemma says that when s and r are chosen "uniformly" under $\mathbf{N}$, then the random trees $\mathcal{T}_{r,\,s}$ and $\mathcal{T}^-_{r,\,s}$ are independent and distributed as $\mathcal{T}_H$ under $\mathbf{N}[\sigma\bullet]$. This result is a consequence of Bismut's decomposition of the excursion of the height process.

Lemma 6.1. *Let H be a continuous height process associated with a general branching mechanism under its excursion measure $\mathbf{N}$. Then for every nonnegative measurable functions f_+ and f_- defined on $\mathcal{C}_+(\mathbb{R}_+)$, we have:*

$$\mathbf{N}\left[\int_0^\sigma \mathrm{d}s \int_0^{H(s)} f_+(H^+_{r,s})\, f_-(H^-_{r,s})\,\mathrm{d}r\right] = \mathbf{N}\left[\sigma f_+(H)\right]\,\mathbf{N}\left[\sigma f_-(H)\right].$$

Remark 6.2. Lemma 6.1 allows to recover directly the distribution of the size of the two fragments given by the fragmentation measure $q^{ske}(\mathrm{d}s,\mathrm{d}r) = 2\beta\sigma_{r,s}^{-1}\mathbf{1}_{[0,H(s)]}(r)\,\mathrm{d}s\,\mathrm{d}r$ on the skeleton in [44, Lemma 5.1]. The Brownian case ($\pi = 0$ and $\beta > 0$) appears already in [8] and then in [3].

Proof. We follow the proof of [17, Lemma 3.4] and use notations from [16] on the càd-làg Markov process process $(\rho_s, \eta_s;\ s \in [0,\sigma])$ under $\mathbf{N}$, which is an $\mathcal{M}(\mathbb{R}_+)^2$-valued process. The process (ρ,η) is a Markov process which allows to recover the (*a priori* non-Markovian) height process as a.s. $[0, H(t)] = \mathrm{Supp}\ (\rho_t) = \mathrm{Supp}\ (\eta_t)$. (The process ρ is called the exploration process associated with H and is strong Markov.) Thanks to [16, Proposition 3.1.3], we have that:

$$\mathbf{N}\left[\int_0^\sigma \mathrm{d}s\, F(\rho_s,\eta_s)\right] = \int \mathbb{M}(\mathrm{d}\mu,\mathrm{d}\nu)\, F(\mu,\nu), \tag{6.3}$$

where $\mathbb{M} = \int_0^\infty \mathrm{d}t\, \mathrm{e}^{-at}\, \mathbb{M}_{[0,t]}$ and, for any interval I, $\mathbb{M}_I$ is the law on $\mathcal{M}(\mathbb{R}_+)^2$ of the pair (μ_I, ν_I) defined by:

$$\begin{aligned}\mu_I(f) &= \int \mathcal{N}(\mathrm{d}r,\mathrm{d}\ell,\mathrm{d}x)\,\mathbf{1}_I(r)\,xf(r) + \beta\int_I \mathrm{d}r\, f(r),\\ \nu_I(f) &= \int \mathcal{N}(\mathrm{d}r,\mathrm{d}\ell,\mathrm{d}x)\,\mathbf{1}_I(r)(\ell - x)f(r) + \beta\int_I \mathrm{d}r\, f(r),\end{aligned}$$

with $\mathcal{N}(\mathrm{d}r,\mathrm{d}\ell,\mathrm{d}x)$ a Poisson point measure on $(\mathbb{R}_+)^3$ with intensity $\mathrm{d}r\,\pi(\mathrm{d}\ell)\,\mathbf{1}_{[0,\ell]}(x)\,\mathrm{d}x$. We write $\tilde\rho = (\rho,\eta)$ and $\tilde\eta = (\eta,\rho)$. We recall that the process $(\rho_s;\ s\in[0,\sigma])$ is strong Markov under $\mathbf{N}$, see [16, Proposition 1.2.3], and the time reversal property of (ρ,η), see [16, Corollary 3.1.6], that is $(\tilde\rho_s;\ s\in[0,\sigma])$ and $\left(\tilde\eta_{(\sigma-s)-};\ s\in[0,\sigma]\right)$ have the same distribution under $\mathbf{N}$.

For a measure μ on $\mathbb{R}_+$ and $u > 0$ we define the measure $\mu^{[u]}$, the measure μ erased up to level u and shifted by u, by $\mu^{[u]}(f) = \int f(r-u)\mathbf{1}_{\{r>u\}}\,\mu(\mathrm{d}r)$ for $f \in \mathcal{B}_+(\mathbb{R}_+)$. We write $\tilde\rho^{[u]} = (\rho^{[u]}, \eta^{[u]})$ and similarly for $\tilde\eta$. Let F_i^ε, for $\varepsilon \in \{+,-\}$ and $i \in \{\mathrm{g},\mathrm{d}\}$, be measurable nonnegative functionals defined on the set of càd-làg $\mathcal{M}(\mathbb{R}_+)^2$-valued functions. We shall compute:

$$\begin{aligned}A = \mathbf{N}\Big[\int_0^\sigma \mathrm{d}s \int_0^{H(s)} \mathrm{d}r\, &F^+_{\mathrm{d}}\left(\tilde\rho^{[r]}_{s+t}; t\in[0,\mathrm{T}^+_{r,s}-s]\right)\right) F^+_{\mathrm{g}}\left(\tilde\eta^{[r]}_{(s-t)-}; t\in[0,\mathrm{T}^-_{r,s}-s]\right)\\ &F^-_{\mathrm{d}}\left(\tilde\rho_{\mathrm{T}^+_{r,s}+t}; t\in[0,\sigma-\mathrm{T}^+_{r,s}]\right) F^-_{\mathrm{g}}\left(\tilde\eta_{(\mathrm{T}^-_{r,s}-t)-}; t\in[0,\mathrm{T}^-_{r,s}]\right)\Big].\end{aligned}$$

We write $\mathbf{1}_{[0,r]}\tilde\rho = (\mathbf{1}_{[0,r]}\rho, \mathbf{1}_{[0,r]}\eta)$. Using the Markov property of $\tilde\rho$ at time s, the time reversal property, again the Markov property of $\tilde\rho$ at time s, (6.3) and the transition kernel of $\tilde\rho$ given in [16, Proposition 3.1.2], we get that:

$$A = \mathbf{N}\left[\int_0^\sigma \mathrm{d}s\int_0^{H(s)} \mathrm{d}r\, G^+\left(\tilde\rho_s^{[r]}\right) G^-\left(\mathbf{1}_{[0,r]}\tilde\rho_s\right)\right],$$

for some measurable nonnegative functions G^- and G^+ such that for $\varepsilon \in \{+,-\}$

$$\mathbb{M}[G^\varepsilon] = \mathbf{N}\left[\int_0^\sigma \mathrm{d}s\, F^\varepsilon_{\mathrm{d}}(\tilde\rho_{s+t}, t\in[0,\sigma-s]) F^\varepsilon_{\mathrm{g}}(\tilde\rho_{(s-t)-}, t\in[0,s])\right]. \tag{6.4}$$

Then using (6.3) and the definition of $\mathbb{M}$, we get, with $\tilde{\mu} = (\mu, \nu)$:

$$\begin{aligned} A &= \int_0^\infty \mathrm{d}t\, \mathrm{e}^{-at} \int_0^t \mathrm{d}r\ \mathbb{M}_{[0,t]}(\mathrm{d}\tilde{\mu})\, G^+\left(\tilde{\mu}^{[r]}\right) G^-\left(\mathbf{1}_{[0,r]}\tilde{\mu}\right) \\ &= \int_0^\infty \mathrm{d}t\, \mathrm{e}^{-at} \int_0^t \mathrm{d}r\ \mathbb{M}_{[0,t-r]}[G^+]\, \mathbb{M}_{[0,r]}[G^-] \\ &= \left(\int_0^\infty \mathrm{d}r\, \mathrm{e}^{-ar}\, \mathbb{M}_{[0,r]}[G^+]\right)\left(\int_0^\infty \mathrm{d}r\, \mathrm{e}^{-ar}\, \mathbb{M}_{[0,r]}[G^-]\right) \\ &= \mathbb{M}[G^+]\, \mathbb{M}[G^-], \end{aligned}$$

where we used the independence property, that is $\mathbb{M}_I * \mathbb{M}_J = \mathbb{M}_{I\cup J}$ when I and J are disjoint, for the second equality. We deduce from (6.4) and the monotone class theorem that for any measurable nonnegative functionals F^+ and F^- defined on the set of càd-làg $\mathcal{M}(\mathbb{R}_+)^2$-valued functions, we have:

$$\begin{aligned} &\mathbf{N}\left[\int_0^\sigma \mathrm{d}s \int_0^{H(s)} \mathrm{d}r\, F^+(\tilde{\rho}_{t+T^-_{r,s}}; t\in[0,\sigma_{r,s}])F^-(\tilde{\rho}_{t+\sigma_{r,s}\mathbf{1}_{\{t>T^-_{r,s}\}}}; t\in[0,\sigma-\sigma_{r,s}])\right] \\ &\qquad = \mathbf{N}\left[\int_0^\sigma \mathrm{d}s\, F^+(\tilde{\rho}_t; t\in[0,\sigma])\right]\mathbf{N}\left[\int_0^\sigma \mathrm{d}s\, F^-(\tilde{\rho}_t; t\in[0,\sigma])\right]. \\ &\qquad = \mathbf{N}\left[\sigma F^+(\tilde{\rho}_t; t\in[0,\sigma])\right]\mathbf{N}\left[\sigma F^-(\tilde{\rho}_t; t\in[0,\sigma])\right]. \end{aligned}$$

Then use that H is a measurable functional of the exploration process $\tilde{\rho}$ to conclude. □

6.2. **First moment of $\Psi_{\mathcal{T}}$.** We start with the main result of this section which gives the first moment of functionals of the stable Lévy tree. Recall that $\mathcal{T}_H$ is the real tree coded by H, see Remark 2.1.

Proposition 6.3. *Let $\mathcal{T}$ be the stable Lévy tree with branching mechanism $\psi(\lambda) = \kappa\lambda^\gamma$ where $\kappa > 0$ and $\gamma \in (1, 2]$. Let $f \in \mathcal{B}_+(\mathbb{T})$, and set $\tilde{f}(T, r) = f(T)$ for $T \in \mathbb{T}$ and $r \in \mathbb{R}_+$. We have:*

$$\mathbb{E}\left[\Psi_{\mathcal{T}}(\tilde{f})\right] = \mathbf{N}\left[\sigma(1-\sigma)^{-1/\gamma} f(\mathcal{T}_H)\mathbf{1}_{\{\sigma<1\}}\right]. \tag{6.5}$$

Proof. Let $f \in \mathcal{B}_+(\mathbb{T})$ and set $\tilde{f}(T, r) = f(T)$ for $T \in \mathbb{T}$ and $r \in \mathbb{R}_+$. Using notations from Section 6.1, we have $\Psi_{\mathcal{T}_H}(\tilde{f}) = \int_0^\sigma \mathrm{d}s \int_0^{H(s)} f(\mathcal{T}_{H^+_{r,s}})\, \mathrm{d}r$. Thus, on the one hand, we get for $\lambda > 0$

$$\begin{aligned} \mathbf{N}\left[\mathrm{e}^{-\lambda\sigma}\, \Psi_{\mathcal{T}_H}(\tilde{f})\right] &= \mathbf{N}\left[\int_0^\sigma \mathrm{d}s \int_0^{H(s)} \mathrm{e}^{-\lambda\sigma^+_{r,s}}\, f(\mathcal{T}_{H^+_{r,s}})\, \mathrm{e}^{-\lambda\sigma^-_{r,s}}\, \mathrm{d}r\right] \\ &= \mathbf{N}\left[\sigma\, \mathrm{e}^{-\lambda\sigma}\right]\mathbf{N}\left[\sigma\, \mathrm{e}^{-\lambda\sigma}\, f(\mathcal{T}_H)\right] \\ &= \mathfrak{g}(0)^2 \int_0^\infty \mathrm{e}^{-\lambda u}\, \mathbf{N}^{(u)}\left[f(\mathcal{T}_H)\right] \frac{\mathrm{d}u}{u^{1/\gamma}} \int_0^\infty \mathrm{e}^{-\lambda y}\, \frac{\mathrm{d}y}{y^{1/\gamma}} \\ &= \mathfrak{g}(0)^2 \int_0^\infty \mathrm{e}^{-\lambda r}\, \mathrm{d}r \int_0^r \mathbf{N}^{(u)}\left[f(\mathcal{T}_H)\right] \frac{\mathrm{d}u}{(u(r-u))^{1/\gamma}}, \end{aligned} \tag{6.6}$$

where we used (6.2) for the first equality, Lemma 6.1 for the second, (4.12) for the third and the change of variable $r = u + y$ for the last. On the other hand, we consider the random variable $H^r = (r^{1-1/\gamma}H(s/r), s \in [0, r])$ for $r > 0$. According to (4.14), H^r under $\mathbf{N}^{(1)}$ is distributed as H

under $\mathbf{N}^{(r)}$. Then, we have for $\lambda > 0$

$$\mathbf{N}\left[\mathrm{e}^{-\lambda\sigma}\,\Psi_{\mathcal{T}_H}(\tilde{f})\right] = \mathfrak{g}(0)\int_0^\infty \mathrm{e}^{-\lambda r}\,\mathbb{E}\left[\Psi_{\mathcal{T}_{H^r}}(\tilde{f})\right]\,\frac{\mathrm{d}r}{r^{1+1/\gamma}}. \tag{6.7}$$

Comparing (6.6) and (6.7), we deduce that $\mathrm{d}r$-a.e., for $r > 0$

$$\mathbb{E}\left[\Psi_{\mathcal{T}_{H^r}}(\tilde{f})\right] = r^{1+1/\gamma}\mathfrak{g}(0)\int_0^r \frac{\mathbf{N}^{(u)}\left[f(\mathcal{T}_H)\right]}{(r-u)^{1/\gamma}}\,\frac{\mathrm{d}u}{u^{1/\gamma}} = r^{1+1/\gamma}\,\mathbf{N}\left[\sigma(r-\sigma)^{-1/\gamma}f(\mathcal{T}_H)\mathbf{1}_{\{\sigma<r\}}\right]. \tag{6.8}$$

From now on, we assume that $f \in \mathcal{C}_+(\mathbb{T})$ is bounded and that there exists $\varepsilon > 0$ such that $f(T) = 0$ if $\mathfrak{m}(T) > 1-\varepsilon$. As $\mathfrak{m}(\mathcal{T}_H) = \sigma$, the map $r \mapsto \mathbf{N}\left[\sigma(r-\sigma)^{-1/\gamma}f(\mathcal{T}_H)\mathbf{1}_{\{\sigma<r\}}\right]$ is continuous at $r = 1$ by dominated convergence. By definition of H^r and the continuity of the height function, we get that a.s. $\lim_{r\to 1}\|H^r - H^1\|_\infty = 0$. Following [2, Proposition 2.10], we get that the $\mathbb{T}$-valued function $r \mapsto \mathcal{T}_{H^r}$ is then a.s. continuous at $r = 1$. We deduce from Proposition 3.3 that $r \mapsto \Psi_{\mathcal{T}_{H^r}}(\tilde{f})$ is continuous at $r = 1$. We also have

$$\Psi_{\mathcal{T}_{H^r}}(\tilde{f}) \leq \mathfrak{m}(\mathcal{T}_{H^r})\mathfrak{h}(\mathcal{T}_{H^r})\,\|f\|_\infty \leq r^{2-1/\gamma}\mathfrak{h}(H^1)\,\|f\|_\infty\,.$$

Since $\mathfrak{h}(H^1)$ is integrable, we deduce by dominated convergence that the map $r \mapsto \mathbb{E}\left[\Psi_{\mathcal{T}_{H^r}}(\tilde{f})\right]$ is continuous at $r = 1$. We deduce from (6.8) that for all $f \in \mathcal{C}_+(\mathbb{T})$ bounded and such that there exists $\varepsilon > 0$ for which $f(T) = 0$ if $\mathfrak{m}(T) > 1-\varepsilon$, we have:

$$\mathbb{E}\left[\Psi_{\mathcal{T}_{H^1}}(\tilde{f})\right] = \mathbf{N}\left[\sigma(1-\sigma)^{-1/\gamma}f(\mathcal{T}_H)\mathbf{1}_{\{\sigma<1\}}\right].$$

By monotone convergence, this equality holds if $f \in \mathcal{C}_+(\mathbb{T})$ is bounded. Then use that $\mathcal{T}_{H^1}$ is distributed as $\mathcal{T}$ to get (6.5). $\square$

The next result is a direct consequence of Proposition 6.3, using that π_*, defined in (4.12), is the distribution of σ under $\mathbf{N}$. Recall the notation $\Psi_{\mathcal{T}}^{\mathfrak{m}\mathfrak{h}}(g(x)h(u))$ which means that g is a function of the mass and h a function of the height.

Corollary 6.4. *Let $\mathcal{T}$ be the stable Lévy tree with branching mechanism $\psi(\lambda) = \kappa\lambda^\gamma$ where $\kappa > 0$ and $\gamma \in (1,2]$. Then we have for every $f \in \mathcal{B}_+([0,1]\times\mathbb{R}_+)$*

$$\mathbb{E}\left[\Psi_{\mathcal{T}}^{\mathfrak{m}\mathfrak{h}}(f)\right] = \mathfrak{g}(0)\int_0^1 x^{-1/\gamma}(1-x)^{-1/\gamma}\,\mathbb{E}\left[f\left(x, x^{1-1/\gamma}\mathfrak{h}(\mathcal{T})\right)\right]\,\mathrm{d}x, \tag{6.9}$$

where $\mathfrak{g}(0)$ is given in (4.3). In particular, we have for every $g \in \mathcal{B}_+([0,1])$

$$\mathbb{E}\left[\Psi_{\mathcal{T}}^{\mathfrak{m}\mathfrak{h}}(g(x))\right] = \mathfrak{g}(0)\int_0^1 x^{-1/\gamma}(1-x)^{-1/\gamma}g(x)\,\mathrm{d}x.$$

Remark 6.5. An equivalent way to state (6.9) is the following equality of measures

$$\mathbb{E}\left[\Psi_{\mathcal{T}}^{\mathfrak{m}\mathfrak{h}}(f)\right] = C(\gamma,\kappa)\,\mathbb{E}\left[f\left(V, V^{1-1/\gamma}\mathfrak{h}(\mathcal{T})\right)\right] \quad \text{with} \quad C(\gamma,\kappa) = \mathrm{B}(1-1/\gamma, 1-1/\gamma)\mathfrak{g}(0),$$

where V is a random variable with distribution $\mathrm{Beta}(1-1/\gamma, 1-1/\gamma)$, independent of $\mathfrak{h}(\mathcal{T})$ and B is the beta function. Using (3.4), this can be interpreted in the following way where we recall that ℓ denotes the length measure on a real tree: taking a stable Lévy tree $\mathcal{T}$ under $\mathbb{P}$ and simultaneously choosing a vertex $y \in \mathcal{T}$ uniformly according to the measure $C(\gamma,\kappa)^{-1}\mu(\mathcal{T}_y)\ell(\mathrm{d}y)$, then the mass and height of the subtree $\mathcal{T}_y$ are jointly distributed as V and $V^{1-1/\gamma}\mathfrak{h}(\mathcal{T})$.

While the measure $\mathbb{E}\left[\Psi_{\mathcal{T}}^{\mathfrak{m}\mathfrak{h}}(\bullet)\right]$ is not known explicitly, its moments can be expressed in terms of the moments of $\mathfrak{h}(\mathcal{T})$.

Corollary 6.6. *Let $\mathcal{T}$ be the stable Lévy tree with branching mechanism $\psi(\lambda) = \kappa\lambda^\gamma$. For every $\alpha, \beta \in \mathbb{C}$ such that $\Re(\gamma\alpha + (\gamma - 1)(\beta + 1)) > 0$, we have*

$$\mathbb{E}\left[\Psi_{\mathcal{T}}^{\mathfrak{m}\mathfrak{h}}(x^\alpha u^\beta)\right] = \mathfrak{g}(0)\mathrm{B}(\alpha + (\beta+1)(1 - 1/\gamma), 1 - 1/\gamma)\,\mathbb{E}\left[\mathfrak{h}(\mathcal{T})^\beta\right], \tag{6.10}$$

where B *is the beta function.*

Observe that $\mathfrak{h}(\mathcal{T})$ has finite moments of all order. This can be seen as a consequence of the convergence in distribution $\frac{b_n}{n}\mathfrak{h}(\tau^n) \xrightarrow{(d)} \mathfrak{h}(\mathcal{T})$ together with the fact that $\left(\frac{b_n}{n}\mathfrak{h}(\tau^n),\, n \in \mathbb{N}\right)$ is bounded in L^p for every $p \in \mathbb{R}$ by Lemma 4.4. The first moment of $\mathfrak{h}(\mathcal{T})$ is given in [18, Proposition 3.4]. We shall discuss the other moments in a future work.

Note that taking $\beta = 0$, we recover [14, Lemma 4.6]. Heuristically, the condition $\Re(\gamma\alpha + (\gamma - 1)(\beta + 1)) > 0$ is due to the fact that under the excursion measure $\mathbf{N}$, the height $\mathfrak{h}$ scales as $\sigma^{1-1/\gamma}$ (see also Lemma 6.11 below), implying that for $\alpha, \beta \in \mathbb{R}$

$$\mathbb{E}\left[\int_{\mathcal{T}} \mu(\mathrm{d}x) \int_0^{H(x)} \mathfrak{m}(\mathcal{T}_{r,x})^\alpha\, \mathfrak{h}(\mathcal{T}_{r,x})^\beta \,\mathrm{d}r\right] < \infty \Longleftrightarrow \mathbb{E}\left[\int_{\mathcal{T}} \mu(\mathrm{d}x) \int_0^{H(x)} \mathfrak{m}(\mathcal{T}_{r,x})^{\alpha+\beta(1-1/\gamma)} \,\mathrm{d}r\right] < \infty.$$

Thus, the condition on α, β corresponds to the phase transition observed in [14, Lemma 4.6 and Remark 4.8] for functionals depending only on the mass (that is $\beta = 0$).

In the Brownian case, $\mathfrak{h}(\mathcal{T})$ is the maximum of the (scaled) Brownian excursion whose moments are known explicitly. Therefore we get an explicit formula for the moments of the measure $\mathbb{E}\left[\Psi_{\mathcal{T}}^{\mathfrak{m}\mathfrak{h}}(\bullet)\right]$.

Corollary 6.7. *Let $\mathcal{T}$ be the Brownian tree with branching mechanism $\psi(\lambda) = \kappa\lambda^2$. For every $\alpha, \beta \in \mathbb{C}$ such that $\Re(2\alpha + \beta + 1) > 0$, we have*

$$\mathbb{E}\left[\Psi_{\mathcal{T}}^{\mathfrak{m}\mathfrak{h}}(x^\alpha u^\beta)\right] = \frac{1}{\sqrt{\pi\kappa}} \left(\frac{\pi}{\kappa}\right)^{\beta/2} \xi(\beta)\mathrm{B}\left(\alpha + \frac{\beta+1}{2}, \frac{1}{2}\right), \tag{6.11}$$

where ξ is the Riemann xi function defined by $\xi(s) = \frac{1}{2}s(s-1)\pi^{-s/2}\Gamma(s/2)\zeta(s)$ for every $s \in \mathbb{C}$ and ζ is the Riemann zeta function.

Proof. The normalized excursion of the height process H is distributed as $\sqrt{2/\kappa}\, B_{\mathrm{ex}}$ where B_{ex} is the normalized Brownian excursion, see *e.g.* [16]. Therefore we get the identity $\mathfrak{h}(\mathcal{T}) \stackrel{(d)}{=} \sqrt{2/\kappa}\, \max B_{\mathrm{ex}}$. By [11, Proposition 2.1 and Eq. (4.10)], we have

$$\mathbb{E}\left[(\max B_{\mathrm{ex}})^\beta\right] = 2\left(\frac{\pi}{2}\right)^{\beta/2} \xi(\beta), \quad \forall \beta \in \mathbb{C}.$$

The result follows then from Corollary 6.6 and the value of $\mathfrak{g}(0)$ given in (4.4) . $\square$

6.3. **Finiteness of $\Psi_{\mathcal{T}}^{\mathfrak{m}\mathfrak{h}}(f)$.** This section is devoted to the study of the finiteness of functionals of the mass and height on the stable Lévy tree. Arguing as in the proof of Lemma 5.2 and using Corollary 6.6 and the fact that $\mathfrak{h}(\mathcal{T})$ has finite moments of all orders, we get the following result.

Lemma 6.8. *Let $\mathcal{T}$ be the stable Lévy tree with branching mechanism $\psi(\lambda) = \kappa\lambda^\gamma$ where $\kappa > 0$ and $\gamma \in (1, 2]$. Let $\alpha, \beta \in \mathbb{R}$ such that $\gamma\alpha + (\gamma - 1)(\beta + 1) > 0$. For and every $p \geq 1$ such that $p(\gamma\alpha + (\gamma - 1)\beta) > 1 - \gamma$ and $\delta \in \mathbb{R}$, we have:*

$$\mathbb{E}\left[\mathfrak{h}(\mathcal{T})^\delta \Psi^{\mathfrak{m}\mathfrak{h}}_{\mathcal{T}}(x^\alpha u^\beta)^p\right] < \infty. \tag{6.12}$$

We now state the main result of this section which gives an integral test for the finiteness of functionals of the mass and height on the stable Lévy tree.

Proposition 6.9. *Let $\mathcal{T}$ be the stable Lévy tree with branching mechanism $\psi(\lambda) = \kappa\lambda^\gamma$ where $\kappa > 0$ and $\gamma \in (1, 2]$. Let $f \in \mathcal{B}_+([0, 1] \times \mathbb{R}_+)$ be of the form $f(x, u) = g(x)u^\beta$ or $f(x, u) = x^\alpha h(u)$ where $\alpha, \beta \in \mathbb{R}$, and g, h nonincreasing. Then we have*

$$\Psi^{\mathfrak{m}\mathfrak{h}}_{\mathcal{T}}(f) \begin{cases} < \infty & a.s., \\ = \infty & a.s., \end{cases} \tag{6.13}$$

according as

$$\int_0 f(x^{\gamma/(\gamma-1)}, x)\,\mathrm{d}x \begin{cases} < \infty, \\ = \infty. \end{cases} \tag{6.14}$$

Furthermore, if $\Psi^{\mathfrak{m}\mathfrak{h}}_{\mathcal{T}}(f)$ is a.s. finite then we have $\mathbb{E}\left[\Psi^{\mathfrak{m}\mathfrak{h}}_{\mathcal{T}}(f)\right] < \infty$.

Proof. We first prove that if $\int_0 f(x^{\gamma/(\gamma-1)}, x)\,\mathrm{d}x$ is finite then $\mathbb{E}\left[\Psi^{\mathfrak{m}\mathfrak{h}}_{\mathcal{T}}(f)\right]$ is finite and thus $\Psi^{\mathfrak{m}\mathfrak{h}}_{\mathcal{T}}(f)$ is a.s. finite.

Let $\beta \in \mathbb{R}$ and $g \in \mathcal{B}_+([0, 1])$ be such that $\int_0 g(x^{\gamma/(\gamma-1)})x^\beta\,\mathrm{d}x < \infty$. Recall that $\mathfrak{h}(\mathcal{T})$ has finite moments of all orders. Thus, by (6.9), we have

$$\mathbb{E}\left[\Psi^{\mathfrak{m}\mathfrak{h}}_{\mathcal{T}}(g(x)u^\beta)\right] = \mathfrak{g}(0)\,\mathbb{E}\left[\mathfrak{h}(\mathcal{T})^\beta\right] \int_0^1 g(x)x^{(\beta+1)(1-1/\gamma)-1}(1 - x)^{-1/\gamma}\,\mathrm{d}x < \infty.$$

Next, let $\alpha \in \mathbb{R}$ and $h \in \mathcal{B}_+(\mathbb{R}_+)$ be nonincreasing such that $\int_0 h(x)x^{\alpha\gamma/(\gamma-1)}\,\mathrm{d}x < \infty$. Again by (6.9), we have

$$\mathbb{E}\left[\Psi^{\mathfrak{m}\mathfrak{h}}_{\mathcal{T}}(x^\alpha h(u))\right] = \mathfrak{g}(0) \int_0^1 x^{\alpha-1/\gamma}(1 - x)^{-1/\gamma}\,\mathbb{E}\left[h\left(x^{1-1/\gamma}\mathfrak{h}(\mathcal{T})\right)\right]\mathrm{d}x.$$

Now, letting k goes to infinity in (5.8) and using the continuity of the cdf of $\mathfrak{h}(\mathcal{T})$ (see [18]), we get that

$$\mathbb{P}(\mathfrak{h}(\mathcal{T}) \leq y) \leq 1 \wedge \left(C_0 \exp\left(-c_0 y^{-\gamma/(\gamma-1)}\right)\right) \quad \text{for all} \quad y \geq 0.$$

We deduce that $\mathfrak{h}(\mathcal{T}) \geq_{\mathrm{st}} Y$ where the cdf of the random variable Y is given by the right-hand side of the inequality above. Using that h is nonincreasing and repeating the same computations as in the proof of Lemma 5.1 (cf. (5.12)), we deduce that

$$\mathbb{E}\left[\Psi^{\mathfrak{m}\mathfrak{h}}_{\mathcal{T}}(x^\alpha h(u))\right] \leq \mathfrak{g}(0) \int_0^1 x^{\alpha-1/\gamma}(1 - x)^{-1/\gamma}\,\mathbb{E}\left[h\left(x^{1-1/\gamma}Y\right)\right]\mathrm{d}x < \infty.$$

This finishes the proof of the finite case. The infinite case is more delicate and its proof is postponed to Section 6.4. □

We end this section with a complete description of the behavior of polynomial functionals of the mass and height on the stable Lévy tree, which is a particular case of Proposition 6.9 (and Lemma 6.8 for $\alpha > 0$ and $\beta > 0$).

Corollary 6.10. *Let $\mathcal{T}$ be the stable Lévy tree with branching mechanism $\psi(\lambda) = \kappa\lambda^\gamma$ with $\kappa > 0$ and $\gamma \in (1, 2]$, and let $\alpha, \beta \in \mathbb{R}$. Then we have*

$$\gamma\alpha + (\gamma - 1)(\beta + 1) > 0 \iff \Psi_{\mathcal{T}}^{\mathfrak{mh}}(x^\alpha u^\beta) < \infty \text{ a.s.} \iff \mathbb{E}\left[\Psi_{\mathcal{T}}^{\mathfrak{mh}}(x^\alpha u^\beta)\right] < \infty, \quad (6.15)$$

$$\gamma\alpha + (\gamma - 1)(\beta + 1) \leq 0 \iff \Psi_{\mathcal{T}}^{\mathfrak{mh}}(x^\alpha u^\beta) = \infty \text{ a.s.} \iff \mathbb{E}\left[\Psi_{\mathcal{T}}^{\mathfrak{mh}}(x^\alpha u^\beta)\right] = \infty. \quad (6.16)$$

6.4. **Proof of the infinite case in Proposition 6.9.** Recall that H denotes the height process under the excursion measure $\mathbf{N}$. Recall that $\sigma_{r,s}$ and $\mathfrak{h}_{r,s}$ are the length and height of the excursion of H above level r that straddles s, see Section 6.1. Let $f \in \mathcal{B}_+([0, 1] \times \mathbb{R}_+)$. Set

$$Z_f = \int_0^\sigma \mathrm{d}s \int_0^{H(s)} f(\sigma_{r,s}, \mathfrak{h}_{r,s})\, \mathrm{d}r. \quad (6.17)$$

Notice that under $\mathbf{N}^{(1)}$, the random variable Z_f is distributed as $\Psi_{\mathcal{T}}^{\mathfrak{mh}}(f)$ under $\mathbb{P}$. Using the scaling property (4.14) of the height process, we have the following more general result which is partially given in [14] (notice that there is a misprint in the first line of p.34 therein).

Lemma 6.11. *Let $\psi(\lambda) = \kappa\lambda^\gamma$ with $\kappa > 0$ and $\gamma \in (1, 2]$ and let H be the ψ-height process. For every $x > 0$, the random variable*

$$\Big((\mathfrak{h}(s),\, s \in [0, x]), (\sigma_{r,s}, H_{r,s};\, r \in [0, H(s)], s \in [0, x])\Big)$$

under $\mathbf{N}^{(x)}$ is distributed as the following random variable under $\mathbf{N}^{(1)}$

$$\left(\left(x^{1-1/\gamma} H(s/x),\, s \in [0, x]\right), \left(x\sigma_{x^{-1+1/\gamma} r, s/x}, x^{1-1/\gamma}\mathfrak{h}_{x^{-1+1/\gamma} r, s/x};\, r \in [0, x^{1-1/\gamma} H(s/x)], s \in [0, x]\right)\right).$$

In particular, the random variable $\Big(\left(H(s),\, s \in [0, x]\right), Z_f\Big)$ under $\mathbf{N}^{(x)}$ is distributed as the random variable $\left(\left(x^{1-1/\gamma} H(s/x),\, s \in [0, x]\right), x^{2-1/\gamma} Z_{f_x}\right)$ under $\mathbf{N}^{(1)}$, where f_x is defined by $f_x(y, u) = f(xy, x^{1-1/\gamma} u)$ for $x > 0$.

Conditionally on H, let U be uniformly distributed on $[0, \sigma]$ under $\mathbf{N}[\sigma\bullet]$. Using Bismut's decomposition, see *e.g.* [17, Theorem 4.5] or [1, Theorem 2.1], we get that under $\mathbf{N}[\sigma\bullet]$, the random variable $H(U)$ has Lebesgue distribution on $(0, \infty)$ and, conditionally on $\{H(U) = t\}$, the process $((\sigma_{t-r,U}, \mathfrak{h}_{t-r,U}),\, 0 \leq r \leq t)$ is distributed as $((\mathsf{S}_r, \mathsf{H}_r),\, 0 \leq r \leq t)$ where

$$\mathsf{S}_r = \sum_{s \leq r} \mathfrak{m}(\mathfrak{T}_s) \quad \text{and} \quad \mathsf{H}_r = \max_{s \leq r} \left(\mathfrak{h}(\mathfrak{T}_s) + r - s\right), \quad \forall 0 \leq r \leq t, \quad (6.18)$$

where $\mathfrak{m}(\mathfrak{T}_s)$ (resp. $\mathfrak{h}(\mathfrak{T}_s)$) stands for the mass (resp. the height) of the real tree $\mathfrak{T}_s$, and $\mathfrak{T} = (\mathfrak{T}_s,\, s \geq 0)$ is a $\mathbb{T}$-valued Poisson point process on $[0, t]$ whose intensity is given below. If $\gamma = 2$, the Poisson point process $\mathfrak{T}$ has intensity $2\kappa\, \mathbf{N}$. To describe the intensity of $\mathfrak{T}$ for $\gamma \in (1, 2)$, we introduce the probability distribution $\mathbf{P}_a$ on $\mathbb{T}$ which is the law of a random tree obtained by gluing a family of trees $(T_i, i \in I)$ at their root, with $\sum_{i\in I} \delta_{T_i}(\mathrm{d}T)$ a $\mathbb{T}$-valued Poisson point measure with intensity $a\, \mathbf{N}[\mathrm{d}T]$, see also [1, Section 2.6] for more details on $\mathbf{P}_a$. If $\gamma \in (1, 2)$, the Poisson point process $\mathfrak{T}$ has intensity $\int_0^\infty a\pi(\mathrm{d}a)\mathbf{P}_a(\mathrm{d}T)$ where π is the Lévy measure associated with ψ given by (4.10). In particular, we get the equality in law

$$\int_0^{H(U)} f(\sigma_{r,U}, \mathfrak{h}_{r,U})\,\mathrm{d}r \text{ under } \mathbf{N}\left[\sigma \bullet | H(U) = t\right] \quad \overset{(d)}{=} \quad \int_0^t f(\mathsf{S}_r, \mathsf{H}_r)\,\mathrm{d}r. \tag{6.19}$$

In the proof of [14, Lemma 4.6], see Section 8.6 and more precisely (8.20) therein, it is proven that S is a stable subordinator with Laplace transform $\mathbb{E}\left[\exp(-\lambda \mathsf{S}_1)\right] = \exp(-\gamma\kappa^{1/\gamma}\lambda^{1-1/\gamma})$. We shall determine the intensity of the Poisson point process $\mathfrak{h}(\mathfrak{T}) = (\mathfrak{h}(\mathfrak{T}_s), 0 \leq s \leq t)$. If $\gamma = 2$, $\mathfrak{h}(\mathfrak{T})$ has intensity $2\kappa\,\mathbf{N}[\mathfrak{h} \in \mathrm{d}x]$. But, by [17, Eq. (14)], we have $\mathbf{N}[\mathfrak{h} > x] = 1/(\kappa x)$. Differentiating with respect to x, we get $\mathbf{N}[\mathfrak{h} \in \mathrm{d}x] = \kappa^{-1}x^{-2}\mathbf{1}_{\{x>0\}}\,\mathrm{d}x$, so that $\mathfrak{h}(\mathfrak{T})$ has intensity $2x^{-2}\mathbf{1}_{\{x>0\}}\,\mathrm{d}x$. If $1 < \gamma < 2$, $\mathfrak{h}(\mathfrak{T})$ has intensity

$$\int_0^\infty a\pi(\mathrm{d}a)\,\mathbf{P}_a(\mathfrak{h} \in \mathrm{d}x).$$

Using (4.13) and the definition of $\mathbf{P}_a$, we have $\mathbf{P}_a(\mathfrak{h} \leq x) = \mathrm{e}^{-a\,\mathbf{N}[\mathfrak{h}>x]} = \mathrm{e}^{-Cax^{-1/(\gamma-1)}}$ where $C = (\kappa(\gamma-1))^{-1/(\gamma-1)}$. Differentiating with respect to x, we obtain

$$\mathbf{P}_a(\mathfrak{h} \in \mathrm{d}x) = \frac{Cax^{-\gamma/(\gamma-1)}}{\gamma-1}\mathrm{e}^{-Cax^{-1/(\gamma-1)}}\mathbf{1}_{\{x>0\}}\,\mathrm{d}x.$$

Since $\pi(\mathrm{d}a) = C'a^{-1-\gamma}\,\mathrm{d}a$ where $C' = \kappa\gamma(\gamma-1)/\Gamma(2-\gamma)$, see (4.10), we deduce that for $x > 0$

$$\begin{aligned}\int_0^\infty a\pi(\mathrm{d}a)\,\mathbf{P}_a(\mathfrak{h} \in \mathrm{d}x) &= \frac{CC'}{\gamma-1}\left(\int_0^\infty a^{1-\gamma}x^{-\gamma/(\gamma-1)}\mathrm{e}^{-Cax^{-1/(\gamma-1)}}\,\mathrm{d}a\right)\mathbf{1}_{\{x>0\}}\,\mathrm{d}x \\ &= \frac{C^{\gamma-1}C'\Gamma(2-\gamma)}{\gamma-1}\mathbf{1}_{\{x>0\}}\,\frac{\mathrm{d}x}{x^2} \\ &= \frac{\gamma}{\gamma-1}\mathbf{1}_{\{x>0\}}\,\frac{\mathrm{d}x}{x^2}.\end{aligned}$$

In all cases, for $\gamma \in (1,2]$, we get that $\mathfrak{h}(\mathfrak{T})$ is a Poisson point process with intensity $(\gamma/(\gamma-1))x^{-2}\mathbf{1}_{\{x>0\}}\,\mathrm{d}x$. Intuitively, this implies that S_r is of order $r^{\gamma/(\gamma-1)}$ while H_r is of order r as $r \to 0$ which, together with (6.19), explains the form of the integral test (6.14).

Our goal now is to show that

$$\int_0 f(x^{\gamma/(\gamma-1)}, x)\,\mathrm{d}x = \infty \quad \Longrightarrow \quad \int_0 f(\mathsf{S}_t, \mathsf{H}_t)\,\mathrm{d}t = \infty \quad \text{a.s.}$$

under the assumptions of Proposition 6.9. To do this, we adapt the proof of Theorem 1 in [19] which gives a necessary and sufficient condition for the divergence of integrals of Lévy processes. We first consider the case $f(x,u) = x^\alpha h(u)$.

Lemma 6.12. *Let $\alpha > -1 + 1/\gamma$ and $h \in \mathcal{B}_+(\mathbb{R}_+)$ be nonincreasing such that $\int_0 h(x)x^{\alpha\gamma/(\gamma-1)}\,\mathrm{d}x = \infty$. We have that a.s.*

$$\int_0 \mathsf{S}_t^\alpha h(\mathsf{H}_t)\,\mathrm{d}t = \infty.$$

Proof. Define the first passage time for $a > 0$

$$\mathrm{T}(a) \coloneqq \inf\left\{t > 0\colon\ \mathsf{H}_t \geq a\right\}. \tag{6.20}$$

Since $t \mapsto \mathsf{H}_t$ is right-continuous, we have

$$\left\{\mathrm{T}(a) > t\right\} = \left\{\mathsf{H}_t < a\right\}. \tag{6.21}$$

Furthermore, since $\mathsf{H}_0 = 0$, it holds that a.s. $\mathrm{T}(a) > 0$ for every $a > 0$.

Set $F(t) = \int_0^t \mathsf{S}_s^\alpha \, \mathrm{d}s$. Clearly $F(t) < \infty$ a.s. if $\alpha \geq 0$. If $-1 + 1/\gamma < \alpha < 0$, we have

$$\mathbb{E}\left[F(t)\right] = \int_0^t \mathbb{E}\left[\mathsf{S}_s^\alpha\right] \, \mathrm{d}s = \mathbb{E}\left[\mathsf{S}_1^\alpha\right] \int_0^t s^{\alpha\gamma/(\gamma-1)} \, \mathrm{d}s,$$

where we used that S is stable with index $1 - 1/\gamma$. Now the last integral is finite because of the condition on α, and

$$\mathbb{E}\left[\mathsf{S}_1^\alpha\right] = \frac{1}{\Gamma(|\alpha|)} \int_0^\infty \mathbb{E}\left[\mathrm{e}^{-\lambda \mathsf{S}_1}\right] \lambda^{-1-\alpha} \, \mathrm{d}\lambda = \frac{1}{\Gamma(|\alpha|)} \int_0^\infty \mathrm{e}^{-\gamma\kappa^{1/\gamma}\lambda^{1-1/\gamma}} \lambda^{-1-\alpha} \, \mathrm{d}\lambda < \infty.$$

Thus, we get $F(t) < \infty$ a.s. for $\alpha > -1 + 1/\gamma$. Furthermore, F is nondecreasing and we have

$$\int_0^1 \mathsf{S}_t^\alpha h(\mathsf{H}_t) \, \mathrm{d}t = \int_0^1 h(\mathsf{H}_t) \, \mathrm{d}F(t). \tag{6.22}$$

We shall need the first and second moment of $F(\mathrm{T}(a))$ for $a > 0$. Using (6.21), we have that

$$\mathbb{E}\left[F(\mathrm{T}(a))\right] = \int_0^\infty \mathbb{E}\left[\mathsf{S}_t^\alpha \mathbf{1}_{\{\mathrm{T}(a) > t\}}\right] \, \mathrm{d}t = \int_0^\infty \mathbb{E}\left[\mathsf{S}_t^\alpha \mathbf{1}_{\{\mathsf{H}_t < a\}}\right] \, \mathrm{d}t.$$

On the other hand, notice that for every $s \in [0, \sigma]$, it holds that $\sigma_{0,s} = \sigma$ is the total mass and $H_{0,s} = \mathfrak{h}$ is the total height. Thus, using Bismut's decomposition, we have

$$\mathbf{N}\left[\sigma^{\alpha+1} \mathbf{1}_{\{\mathfrak{h} < a\}}\right] = \int_0^\infty \mathbf{N}\left[\sigma \sigma_{0,U}^\alpha \mathbf{1}_{\{H_{0,U} < a\}} \middle| H(U) = t\right] \, \mathrm{d}t = \int_0^\infty \mathbb{E}\left[\mathsf{S}_t^\alpha \mathbf{1}_{\{\mathsf{H}_t < a\}}\right] \, \mathrm{d}t, \tag{6.23}$$

where we recall that conditionally on H, under $\mathbf{N}[\sigma \bullet]$, U is uniformly distributed on $[0, \sigma]$ and $(\sigma_{0,U}, H_{0,U})$ conditionally on $\{H(U) = t\}$ is then distributed as $(\mathsf{S}_t, \mathsf{H}_t)$. We deduce that

$$\begin{aligned}
\mathbb{E}\left[F(\mathrm{T}(a))\right] &= \mathbf{N}\left[\sigma^{\alpha+1} \mathbf{1}_{\{\mathfrak{h} < a\}}\right] \\
&= \mathfrak{g}(0) \int_0^\infty x^{-1-1/\gamma} \, \mathbf{N}^{(x)}\left[\sigma^{\alpha+1} \mathbf{1}_{\{\mathfrak{h} < a\}}\right] \, \mathrm{d}x \\
&= \mathfrak{g}(0) \int_0^\infty x^{\alpha-1/\gamma} \, \mathbf{N}^{(1)}\left[x^{1-1/\gamma} \mathfrak{h} < a\right] \, \mathrm{d}x \\
&= \frac{\gamma \mathfrak{g}(0)}{(\alpha+1)\gamma - 1} \, \mathbf{N}^{(1)}\left[\mathfrak{h}^{-1-\alpha\gamma/(\gamma-1)}\right] a^{1+\alpha\gamma/(\gamma-1)},
\end{aligned} \tag{6.24}$$

where we disintegrated with respect to σ for the second equality and used the scaling property (4.14) of the height process for the third. Recall that $\mathfrak{h}$ has finite moments of all orders under $\mathbf{N}^{(1)}$, so that $\mathbb{E}[F(\mathrm{T}(a))]$ is finite for all $a > 0$. Next, set

$$Z_\alpha^{\mathfrak{m}} = \int_0^\sigma \mathrm{d}s \int_0^{H(s)} \sigma_{r,s}^\alpha \, \mathrm{d}r.$$

It follows from Lemma 6.11 that under $\mathbf{N}^{(x)}$, $(\mathfrak{h}, Z_\alpha^{\mathfrak{m}})$ is distributed as $(x^{1-1/\gamma}\mathfrak{h}, x^{\alpha+2-1/\gamma} Z_\alpha^{\mathfrak{m}})$ under $\mathbf{N}^{(1)}$. Recall that $\alpha > -1 + 1/\gamma$. Thus, using Bismut's decomposition as in (6.23), we have

$$\begin{aligned}
\mathbb{E}\left[F(\mathrm{T}(a))^2\right] &= 2\,\mathbb{E}\left[\int_0^\infty \mathsf{S}_t^\alpha \mathbf{1}_{\{\mathsf{H}_t < a\}} \, \mathrm{d}t \int_0^t \mathsf{S}_s^\alpha \, \mathrm{d}s\right] \\
&= 2\,\mathbf{N}\left[\sigma^{\alpha+1} \mathbf{1}_{\{\mathfrak{h} < a\}} \int_0^{H(U)} \sigma_{r,U}^\alpha \, \mathrm{d}r\right] \\
&= 2\,\mathbf{N}\left[\sigma^\alpha \mathbf{1}_{\{\mathfrak{h} < a\}} Z_\alpha^{\mathfrak{m}}\right] \\
&= 2\mathfrak{g}(0) \int_0^\infty x^{-1-1/\gamma} \, \mathbf{N}^{(x)}\left[\sigma^\alpha \mathbf{1}_{\{\mathfrak{h} < a\}} Z_\alpha^{\mathfrak{m}}\right] \, \mathrm{d}x
\end{aligned}$$

$$= 2\mathfrak{g}(0)\int_0^\infty x^{-1-1/\gamma}\,\mathbf{N}^{(1)}\left[x^\alpha \mathbf{1}_{\{x^{1-1/\gamma}\mathfrak{h}<a\}}x^{\alpha+2-1/\gamma}Z_\alpha^{\mathfrak{m}}\right]\,\mathrm{d}x$$

$$= \frac{\mathfrak{g}(0)}{\alpha+1-1/\gamma}\,\mathbf{N}^{(1)}\left[\mathfrak{h}^{-2(1+\alpha\gamma/(\gamma-1))}Z_\alpha^{\mathfrak{m}}\right]a^{2(1+\alpha\gamma/(\gamma-1))}, \tag{6.25}$$

where the last term is finite by (6.12). Combining (6.24) and (6.25) and using Cauchy-Schwartz inequality, we deduce that there exists some finite constant $C > 0$ such that for all $a, b > 0$

$$\mathbb{E}\left[F(\mathrm{T}(a))F(\mathrm{T}(b))\right] \le \mathbb{E}\left[F(\mathrm{T}(a))^2\right]^{1/2}\mathbb{E}\left[F(\mathrm{T}(b))^2\right]^{1/2} \le C\,\mathbb{E}\left[F(\mathrm{T}(a))\right]\mathbb{E}\left[F(\mathrm{T}(b))\right]. \tag{6.26}$$

For $i \in \mathbb{N}$, put $\mathrm{T}_i = \mathrm{T}(2^{-i})$, $h_i = h(2^{-i})$ and $\Delta h_i = h_{i+1} - h_i$. Notice that the sequence $(\mathrm{T}_i,\, i \in \mathbb{N})$ is nonincreasing and $\Delta h_i \ge 0$. Set $V_n = \sum_{i=1}^n F(\mathrm{T}_i)\Delta h_{i-1}$. Notice that $\mathbb{E}[V_n]$ is finite as $\mathbb{E}[F(\mathrm{T}(a))]$ is finite for all $a > 0$. By (6.26), we have

$$\begin{aligned}
\mathbb{E}\left[V_n^2\right] &= \sum_{i=1}^n \mathbb{E}\left[F(\mathrm{T}_i)^2\right](\Delta h_{i-1})^2 + 2\sum_{1\le i<j\le n}\mathbb{E}\left[F(\mathrm{T}_i)F(\mathrm{T}_j)\right]\Delta h_{i-1}\Delta h_{j-1}\\
&\le C\sum_{i=1}^n \mathbb{E}\left[F(\mathrm{T}_i)\right]^2(\Delta h_{i-1})^2 + 2C\sum_{1\le i<j\le n}\mathbb{E}\left[F(\mathrm{T}_i)\right]\mathbb{E}\left[F(\mathrm{T}_j)\right]\Delta h_{i-1}\Delta h_{j-1}\\
&= C\left(\sum_{i=1}^n \mathbb{E}\left[F(\mathrm{T}_i)\right]\Delta h_{i-1}\right)^2 = C\,\mathbb{E}\left[V_n\right]^2.
\end{aligned}$$

Therefore, we get that $\limsup_n \mathbb{E}\left[V_n\right]^2/\mathbb{E}\left[V_n^2\right] > 0$. By [32], it follows that

$$\mathbb{P}\left(\limsup_n \frac{V_n}{\mathbb{E}\left[V_n\right]} \ge 1\right) > 0. \tag{6.27}$$

Using (6.24), notice that for some finite constant $C > 0$, we have

$$\begin{aligned}
\int_0^1 x^{1+\alpha\gamma/(\gamma-1)}\,|\mathrm{d}h(x)| &\le \sum_{i=1}^\infty (2^{-i+1})^{1+\alpha\gamma/(\gamma-1)}\int_{2^{-i}}^{2^{-i+1}}|\mathrm{d}h(x)|\\
&= C\sum_{i=1}^\infty \mathbb{E}\left[F(\mathrm{T}_i)\right]\Delta h_{i-1} = C\lim_{n\to\infty}\mathbb{E}\left[V_n\right].
\end{aligned} \tag{6.28}$$

Since $\int_0^1 x^{1+\alpha\gamma/(\gamma-1)}\,|\mathrm{d}h(x)| \ge -h(1) + (1+\alpha\gamma/(\gamma-1))\int_0^1 h(x)x^{\alpha\gamma/(\gamma-1)}\,\mathrm{d}x = \infty$ by assumption, it follows from (6.28) that $\lim_{n\to\infty}\mathbb{E}\left[V_n\right] = \infty$. Thus, using (6.27) and the fact that V_n is nondecreasing, we deduce that $\lim_{n\to\infty} V_n = \infty$ with positive probability, that is

$$\mathbb{P}\left(\sum_{i=1}^\infty F(\mathrm{T}_i)\Delta h_{i-1} = \infty\right) > 0. \tag{6.29}$$

Since h is nonincreasing, we have

$$\int_0^{\mathrm{T}_0} h(\mathsf{H}_t)\,\mathrm{d}F(t) \ge \sum_{i=0}^\infty h_{i-1}\left(F(\mathrm{T}_{i-1}) - F(\mathrm{T}_i)\right). \tag{6.30}$$

A summation by parts gives

$$\sum_{i=1}^n h_{i-1}\left(F(\mathrm{T}_{i-1}) - F(\mathrm{T}_i)\right) = F(\mathrm{T}_0)h_0 - F(\mathrm{T}_n)h_n + \sum_{i=1}^n F(\mathrm{T}_i)\Delta h_{i-1}. \tag{6.31}$$

But, notice that

$$F(\mathrm{T}_n)h_n = F(\mathrm{T}_n)h(2^{-n}) \leq \int_0^{\mathrm{T}_n} h(\mathsf{H}_t)\,\mathrm{d}F(t) \leq \int_0^{\mathrm{T}_0} h(\mathsf{H}_t)\,\mathrm{d}F(t).$$

Together with (6.30) and (6.31), this yields

$$F(\mathrm{T}_0)h_0 + \sum_{i=1}^{\infty} F(\mathrm{T}_i)\Delta h_{i-1} \leq 2\int_0^{\mathrm{T}_0} h(\mathsf{H}_t)\,\mathrm{d}F(t).$$

It follows from (6.29) that $\int_0^{\mathrm{T}_0} \mathsf{S}_t^{\alpha} h(\mathsf{H}_t)\,\mathrm{d}t = \int_0^{\mathrm{T}_0} h(\mathsf{H}_t)\,\mathrm{d}F(t)$ diverges with positive probability.

Finally, since the event $\{\int_0 \mathsf{S}_t^{\alpha} h(\mathsf{H}_t)\,\mathrm{d}t = \infty\}$ is $\mathcal{F}_{0+}$-measurable where $(\mathcal{F}_t)_{t\geq 0}$ is the filtration generated by the Poisson point process $\mathfrak{T}$, Blumenthal's zero-one law entails that $\int_0^1 \mathsf{S}_t^{\alpha} h(\mathsf{H}_t)\,\mathrm{d}t$ diverges with probability 1. □

Lemma 6.13. *Let $\beta > -1$ and $g \in \mathcal{B}_+([0,1])$ be nonincreasing such that $\int_0 g(x^{\gamma/(\gamma-1)})x^{\beta}\,\mathrm{d}x = \infty$. We have that a.s.*

$$\int_0 g(\mathsf{S}_t)\mathsf{H}_t^{\beta}\,\mathrm{d}t = \infty.$$

Proof. The proof is similar to that of Lemma 6.12 and we only highlight the major differences. Define the first passage time $\mathrm{T}(a) = \inf\{t > 0\colon \mathsf{S}_t > a\}$ for every $a > 0$. Since S is a stable subordinator, we have a.s. $\mathrm{T}(a) > 0$ for every $a > 0$. Set $F(t) = \int_0^t \mathsf{H}_s^{\beta}\,\mathrm{d}s$. Notice that $F(t) < \infty$ a.s. if $\beta \geq 0$. If $-1 < \beta < 0$, then using that $\mathsf{H}_s \geq s$, we have a.s. $F(t) \leq \int_0^t s^{\beta}\,\mathrm{d}s < \infty$. To compute the first moment of $F(\mathrm{T}(a))$, use Bismut's decomposition as in (6.23) to get

$$\begin{aligned}
\mathbb{E}\left[F(\mathrm{T}(a))\right] &= \mathbb{E}\left[\int_0^{\infty} \mathsf{H}_t^{\beta}\mathbf{1}_{\{\mathsf{S}_t<a\}}\,\mathrm{d}t\right] \\
&= \mathbf{N}\left[\sigma\mathbf{1}_{\{\sigma<a\}}\mathfrak{h}^{\beta}\right] \\
&= \mathfrak{g}(0)\int_0^a x^{(\beta+1)(1-1/\gamma)-1}\,\mathbf{N}^{(1)}\left[\mathfrak{h}^{\beta}\right]\mathrm{d}x \\
&= \frac{\mathfrak{g}(0)}{(\beta+1)(1-1/\gamma)}\,\mathbf{N}^{(1)}\left[\mathfrak{h}^{\beta}\right]a^{(\beta+1)(1-1/\gamma)}.
\end{aligned} \tag{6.32}$$

Setting

$$Z_{\beta}^{\mathfrak{h}} = \int_0^{\sigma}\mathrm{d}s\int_0^{H(s)} H_{r,s}^{\beta}\,\mathrm{d}r$$

and using Bismut's decomposition as in (6.23) and the fact that under $\mathbf{N}^{(x)}$, $(\mathfrak{h}, Z_{\beta}^{\mathfrak{h}})$ is distributed as $(x^{1-1/\gamma}\mathfrak{h}, x^{(\beta+1)(1-1/\gamma)+1}Z_{\beta}^{\mathfrak{h}})$ under $\mathbf{N}^{(1)}$ by Lemma 6.11, we have

$$\begin{aligned}
\mathbb{E}\left[F(\mathrm{T}(a))^2\right] &= 2\,\mathbb{E}\left[\int_0^{\infty} \mathsf{H}_t^{\beta}\mathbf{1}_{\{\mathsf{S}_t<a\}}\,\mathrm{d}t\int_0^t \mathsf{H}_s^{\beta}\,\mathrm{d}s\right] \\
&= 2\,\mathbf{N}\left[\sigma\mathbf{1}_{\{\sigma<a\}}\mathfrak{h}^{\beta}\int_0^{H(U)} H_{r,U}^{\beta}\,\mathrm{d}r\right] \\
&= 2\,\mathbf{N}\left[\mathbf{1}_{\{\sigma<a\}}\mathfrak{h}^{\beta}Z_{\beta}^{\mathfrak{h}}\right] \\
&= 2\mathfrak{g}(0)\int_0^a x^{-1-1/\gamma}\,\mathbf{N}^{(x)}\left[\mathfrak{h}^{\beta}Z_{\beta}^{\mathfrak{h}}\right]\mathrm{d}x \\
&= \frac{\mathfrak{g}(0)}{(\beta+1)(1-1/\gamma)}\,\mathbf{N}^{(1)}\left[\mathfrak{h}^{\beta}Z_{\beta}^{\mathfrak{h}}\right]a^{2(\beta+1)(1-1/\gamma)},
\end{aligned} \tag{6.33}$$

where $\mathbf{N}^{(1)}\left[\mathfrak{h}^{\beta} Z^{\mathfrak{h}}_{\beta}\right] < \infty$ by (6.12). Combining (6.32) and (6.33), we see that the estimate (6.26) holds. The rest of the proof is similar to that of Lemma 6.12 (with h_i replaced by $g_i = g(2^{-i})$). $\square$

We can now finish the proof of Proposition 6.9. Let $f \in \mathcal{B}_+([0,1]\times\mathbb{R}_+)$ be of the form $f(x,u) = g(x)u^{\beta}$ or $f(x,u) = x^{\alpha}h(u)$ with g, h nonincreasing and such that $\int_0 f(x^{\gamma/(\gamma-1)}, x)\,\mathrm{d}x = \infty$. By Lemmas 6.12 and 6.13, we have that, in the cases $\alpha > -1 + 1/\gamma$ and $\beta > -1$, a.s.

$$\int_0 f(\mathsf{S}_t, \mathsf{H}_t)\,\mathrm{d}t = \infty. \tag{6.34}$$

Now suppose that $\alpha \le -1 + 1/\gamma$. Since h is nonincreasing and satisfies $\int_0 h(x)x^{\alpha\gamma/(\gamma-1)}\,\mathrm{d}x = \infty$, there exists a constant $C > 0$ such that $h \ge C$ on some interval $(0,\varepsilon)$. Thus, we have

$$\int_0 \mathsf{S}_t^{\alpha} h(\mathsf{H}_t)\,\mathrm{d}t \ge C \int_0 \mathsf{S}_t^{\alpha}\,\mathrm{d}t,$$

where the last integral diverges a.s. by Lemma 6.13 as $\int_0 x^{\alpha\gamma/(\gamma-1)}\,\mathrm{d}x = \infty$. Similarly, if $\beta \le -1$, there exists a constant $C' > 0$ such that $g \ge C'$ on $(0,\varepsilon)$. Thus, we have

$$\int_0 g(\mathsf{S}_t)\mathsf{H}_t^{\beta}\,\mathrm{d}t \ge C' \int_0 \mathsf{H}_t^{\beta}\,\mathrm{d}t,$$

and the last integral diverges by Lemma 6.12 since $\int_0 x^{\beta}\,\mathrm{d}x = \infty$. This proves that (6.34) holds for all $\alpha, \beta \in \mathbb{R}$.

Combining (6.19) and (6.34), we deduce that

$$\begin{aligned}
\mathbf{N}\left[\sigma;\, Z_f < \infty\right] &= \mathbf{N}\left[\sigma;\, \sigma\int_0^{H(U)} f(\sigma_{r,U}, H_{r,U})\,\mathrm{d}r < \infty\right] \\
&= \int_0^{\infty} \mathbf{N}\left[\sigma;\, \sigma\int_0^{H(U)} f(\sigma_{r,U}, H_{r,U})\,\mathrm{d}r < \infty \middle| H(U) = t\right]\,\mathrm{d}t \\
&= \int_0^{\infty} \mathbb{P}\left(\mathsf{S}_t \int_0^t f(\mathsf{S}_r, \mathsf{H}_r)\,\mathrm{d}r < \infty\right)\,\mathrm{d}t = 0.
\end{aligned}$$

It follows that $\mathbf{N}$-a.e. $Z_f = \infty$. Disintegrating with respect to σ and using the scaling property from Lemma 6.11, we get

$$0 = \mathbf{N}\left[Z_f < \infty\right] = \int_0^{\infty} \mathbf{N}^{(x)}\left[Z_f < \infty\right]\,\pi_*(\mathrm{d}x) = \int_0^{\infty} \mathbf{N}^{(1)}\left[x^{2-1/\gamma} Z_{f_x} < \infty\right]\,\pi_*(\mathrm{d}x).$$

Consequently, $\mathrm{d}x$-a.e. on $(0,\infty)$, we have $\mathbf{N}^{(1)}\left[Z_{f_x} < \infty\right] = 0$. Suppose that $f(y,u) = g(y)u^{\beta}$ with g nonincreasing. Then, under $\mathbf{N}^{(1)}$, Z_{f_x} is equal to $x^{\beta(1-1/\gamma)}\int_0^1 \mathrm{d}s \int_0^{H(s)} g(x\sigma_{r,s})H_{r,s}^{\beta}\,\mathrm{d}r$ and we get that

$$x \mapsto \mathbf{N}^{(1)}\left[\int_0^1 \mathrm{d}s \int_0^{H(s)} g(x\sigma_{r,s})H_{r,s}^{\beta}\,\mathrm{d}r < \infty\right]$$

vanishes $\mathrm{d}x$-a.e. on $(0,\infty)$. Moreover, this function is nonincreasing in x as g is nonincreasing. Hence it is identically zero. In particular, taking $x = 1$ yields $\mathbf{N}^{(1)}\left[Z_f < \infty\right] = 0$, and thus $\Psi_{\mathcal{T}}^{\mathfrak{m}\mathfrak{h}}(f) = +\infty$ a.s. as Z_f under $\mathbf{N}^{(1)}$ is distributed as $\Psi_{\mathcal{T}}^{\mathfrak{m}\mathfrak{h}}(f)$. The same argument applies if we suppose that $f(y,u) = y^{\alpha}h(u)$ instead. This completes the proof.

7. Phase transition for functionals of the mass and height

Recall that τ^n is a BGW(ξ) conditioned to have n vertices (with $n \in \Delta$) and ξ satisfies (ξ1) and (ξ2)$'$, with the sequence $(b_n, n \in \mathbb{N}^*)$ in (4.1), and that $\mathcal{T}$ is a stable Lévy tree with branching mechanism $\psi(\lambda) = \kappa\lambda^\gamma$. In this section, we study the limit of

$$\mathcal{A}_n^\circ(f) = \frac{b_n}{n^2} \sum_{w \in \tau^{n,\circ}} |\tau_w^n| f\left(\frac{b_n}{n}\tau_w^n, \frac{b_n}{n} H(w)\right)$$

for functions $f \in \mathcal{B}(\mathbb{T} \times \mathbb{R}_+)$ continuous on $(\mathbb{T} \setminus \mathbb{T}_0) \times \mathbb{R}_+$ but that may blow up as either the mass or the height goes to 0.

7.1. **A general convergence result.**

We now give a first convergence result for general functionals that may blow up. Recall from (2.5) the definition of $\mathbb{T}_0$. Notice that $\mathcal{A}_n^\circ(\mathbb{T}_0 \times \mathbb{R}_+) = 0$ and $\Psi_\mathcal{T}(\mathbb{T}_0 \times \mathbb{R}_+) = 0$.

Proposition 7.1. *Assume that ξ satisfies (ξ1) and (ξ2)$'$. Let $f \in \mathcal{B}(\mathbb{T} \times \mathbb{R}_+)$ be continuous on $(\mathbb{T} \setminus \mathbb{T}_0) \times \mathbb{R}_+$ and $\alpha, \beta \in \mathbb{R}$ with $\gamma\alpha + (\gamma - 1)(\beta + 1) > 0$ be such that*

$$|f(T, r)| \le C\, \mathfrak{m}(T)^\alpha \mathfrak{h}(T)^\beta, \quad \text{for all } T \in \mathbb{T} \setminus \mathbb{T}_0 \text{ and } r \ge 0, \tag{7.1}$$

for some finite constant $C > 0$. Then $\Psi_\mathcal{T}(|f|)$ is a.s. finite and we have the convergence in distribution

$$\mathcal{A}_n^\circ(f) = \frac{b_n}{n^2} \sum_{w \in \tau^{n,\circ}} |\tau_w^n| f\left(\frac{b_n}{n}\tau_w^n, \frac{b_n}{n} H(w)\right) \xrightarrow{(d)} \Psi_\mathcal{T}(f). \tag{7.2}$$

We also have the convergence of all moments of order $p \ge 1$ such that $p(\gamma\alpha + (\gamma - 1)\beta) > 1 - \gamma$.

Proof. By Corollary 4.10, we know that $\mathcal{A}_n^\circ \xrightarrow{(d)} \Psi_\mathcal{T}$ in the space $\mathcal{M}(\mathbb{T} \times \mathbb{R}_+)$. In particular, the sequence $(\mathcal{A}_n^\circ,\, n \in \Delta)$ is tight (in distribution) in $\mathcal{M}(\mathbb{T} \times \mathbb{R}_+)$, and applying [30, Theorem 4.10], we have

$$\inf_{K \in \mathcal{K}} \sup_{n \in \Delta} \mathbb{E}\left[1 \wedge \mathcal{A}_n^\circ(K^c)\right] = 0, \tag{7.3}$$

where $\mathcal{K}$ is the set of all compact subsets of $\mathbb{T} \times \mathbb{R}_+$. We start by showing that

$$\inf_{K \in \mathcal{K}} \sup_{n \in \Delta} \mathbb{E}\left[\mathcal{A}_n^\circ(K^c)\right] = 0. \tag{7.4}$$

Let $K \in \mathcal{K}$. Using the inequality $x \le 1 \wedge x + x\sqrt{1 \wedge x}$ with $x = \mathcal{A}_n^\circ(K^c) \ge 0$ and the Cauchy–Schwartz inequality, we get that

$$\mathbb{E}\left[\mathcal{A}_n^\circ(K^c)\right] \le \mathbb{E}\left[1 \wedge \mathcal{A}_n^\circ(K^c)\right] + \sqrt{\mathbb{E}\left[\mathcal{A}_n^\circ(1)^2\right] \mathbb{E}\left[1 \wedge \mathcal{A}_n^\circ(K^c)\right]}. \tag{7.5}$$

Since $\mathcal{A}_n^\circ(1) \le \frac{b_n}{n}\mathfrak{h}(\tau^n)$ by (4.21), Lemma 4.4 implies that

$$\sup_{n \in \Delta} \mathbb{E}\left[\mathcal{A}_n^\circ(1)^2\right]^{1/2} \le \sup_{n \in \Delta} \mathbb{E}\left[\left(\frac{b_n}{n}\mathfrak{h}(\tau^n)\right)^2\right]^{1/2} < \infty.$$

This, in conjunction with (7.3) and (7.5), proves (7.4).

Let $\alpha, \beta \in \mathbb{R}$ such that $\gamma\alpha + (\gamma - 1)(\beta + 1) > 0$. We consider the space $S = \mathbb{T} \times \mathbb{R}_+$ with the metric $\rho((T, r), (T', r')) = d_{\mathrm{GHP}}(T, T') + |r - r'|$ and $S_0 = \mathbb{T}_0 \times \mathbb{R}_+$, so that (S, ρ) is a Polish metric space and S_0 is a closed subset of S. We shall consider $0_S = (\{\emptyset\}, 0) \in S_0$ as a distinguished point.

We shall construct a family of functions $\mathfrak{F}$ on S satisfying assumptions (H1)–(H4) of Appendix A in order to apply Proposition A.10. Let $(\delta_k,\, k \in \mathbb{N})$ be a positive increasing sequence such that $(2\gamma-1)\delta_k < (\gamma-1) + (\gamma\alpha + (\gamma-1)\beta) \wedge 0$ for all $k \in \mathbb{N}$. Define for every $k \in \mathbb{N}$

$$f_k(T,r) = \left(\mathfrak{m}(T)^{\delta_k} \vee \mathfrak{m}(T)^{-\delta_k}\right)\left(\mathfrak{h}(T)^{\delta_k} \vee \mathfrak{h}(T)^{-\delta_k}\right) \quad \text{and} \quad g_k(T,r) = \mathfrak{m}(T)^{\alpha}\mathfrak{h}(T)^{\beta} f_k(T,r),$$

for all $T \in \mathbb{T} \setminus \mathbb{T}_0$ and $r \geq 0$ and $f_k = g_k = +\infty$ on $\mathbb{T}_0 \times \mathbb{R}_+$. The functions f_k and g_k are positive and continuous on $(\mathbb{T} \setminus \mathbb{T}_0) \times \mathbb{R}_+$. We define $\mathfrak{F} = \{\mathbf{1}\} \cup \{f_k, g_k \colon k \in \mathbb{N}\}$. Therefore assumptions (H1) and (H2) are satisfied. Notice that $\rho((T,r), S_0) = d_{\mathrm{GHP}}(T, \mathbb{T}_0)$. Let $\varepsilon > 0$ and $M > 0$. By (2.4), $d_{\mathrm{GHP}}(T, \{\emptyset\}) \leq M$ implies that $\mathfrak{h}(T) \leq 2M$ and $\mathfrak{m}(T) \leq M$. Similarly, by Lemma 2.2, $d_{\mathrm{GHP}}(T, \mathbb{T}_0) \geq \varepsilon$ implies that $\mathfrak{h}(T) \geq \varepsilon$ and $\mathfrak{m}(T) \geq \varepsilon$. Therefore, we have the inclusion

$$\{(T,r) \in S \colon \rho((T,r), S_0) \geq \varepsilon,\, \rho((T,r), 0_S) \leq M\} \subset \{T \in \mathbb{T} \colon \mathfrak{h}(T) \in [\varepsilon, 2M],\, \mathfrak{m}(T) \in [\varepsilon, M]\} \times \mathbb{R}_+.$$

Since f_k and g_k are clearly bounded away from zero and infinity on the latter set, assumption (H3) is satisfied. Moreover, f_k/f_{k+1} and g_k/g_{k+1} are continuous and bounded on $S_0^c = (\mathbb{T} \setminus \mathbb{T}_0) \times \mathbb{R}_+$ for every $k \in \mathbb{N}$. Recall that $\rho((T,r), S_0) = d_{\mathrm{GHP}}(T, \mathbb{T}_0)$. Therefore, as $\rho((T,r), S_0) \to 0$, we have $\mathfrak{h}(T) \wedge \mathfrak{m}(T) \to 0$ by Lemma 2.2. It follows that $f_k(T,r)/f_{k+1}(T,r) \to 0$ and $g_k(T,r)/g_{k+1}(T,r) \to 0$ as $\rho((T,r), S_0) \to 0+$. Recall the notation $\mathfrak{F}^\star(f)$ from (H4). We deduce that $f_{k+1} \in \mathfrak{F}^\star(f_k)$ and $g_{k+1} \in \mathfrak{F}^\star(g_k)$ for $k \in \mathbb{N}^*$. We also have that $1/f_1$ is continuous and bounded on S_0^c and that $1/f_1(T,r) \to 0$ as $\rho((T,r), S_0) \to 0+$. This implies that $f_1 \in \mathfrak{F}^\star(\mathbf{1})$. Therefore, assumption (H4) is satisfied.

In order to apply Proposition A.10 to the sequence of measures $(\mathcal{A}_n^\circ,\, n \in \Delta)$ and the family $\mathfrak{F}$, we shall check that the sequence $(\mathcal{A}_n^\circ,\, n \in \Delta)$ is tight (in distribution) in the space $\mathcal{M}_{\mathfrak{F}}$ (see Appendix A for the definition of $\mathcal{M}_{\mathfrak{F}}$). Thanks to Proposition A.4, the sequence $(\mathcal{A}_n^\circ,\, n \in \Delta)$ is tight in the space $\mathcal{M}_{\mathfrak{F}}$ if and only if $(\mathrm{f}\mathcal{A}_n^\circ,\, n \in \Delta)$ is tight in $\mathcal{M}(S)$ for all $\mathrm{f} \in \mathfrak{F}$. Let $\mathrm{f} \in \mathfrak{F}$. Notice that for every $T \in \mathbb{T} \setminus \mathbb{T}_0$ and $r \geq 0$, we have

$$\mathrm{f}((T,r)) \leq \sum_{1 \leq i,j \leq 2} \mathfrak{m}(T)^{\alpha_i}\mathfrak{h}(T)^{\beta_j}$$

for $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbb{R}$ such that $\gamma\alpha_i + (\gamma-1)(\beta_j+1) > 0$ holds for every $i,j \in \{1,2\}$. Therefore, by Lemma 5.2, we have for some $p > 1$ small enough

$$\sup_{n \in \Delta} \mathbb{E}\left[\mathcal{A}_n^\circ(\mathrm{f})^p\right] < \infty \quad \text{and} \quad \sup_{n \in \Delta} \mathbb{E}\left[\mathcal{A}_n^\circ(\mathrm{f}^p)\right] < \infty. \tag{7.6}$$

The first bound gives that (A.3) holds for all $\mathrm{f} \in \mathfrak{F}$ by the Markov inequality. Recall that $\mathcal{K}$ denotes the set of compact subsets of $\mathbb{T} \times \mathbb{R}_+$. Moreover, with q such that $1/p + 1/q = 1$ and $K \in \mathcal{K}$, using Hölder's inequality, we get

$$\mathbb{E}\left[\mathcal{A}_n^\circ(\mathrm{f}\mathbf{1}_{K^c})\right] \leq \mathbb{E}\left[\mathcal{A}_n^\circ(\mathbf{1}_{K^c})\right]^{1/q} \mathbb{E}\left[\mathcal{A}_n^\circ(\mathrm{f}^p)\right]^{1/p}.$$

Using the second bound in (7.6) and (7.4), we deduce that

$$\inf_{K \in \mathcal{K}} \sup_{n \in \Delta} \mathbb{E}\left[\mathcal{A}_n^\circ(\mathrm{f}\mathbf{1}_{K^c})\right] = 0.$$

Thus (A.4) holds for all $\mathrm{f} \in \mathfrak{F}$. According to Proposition A.4-(i), we get that the sequence $(\mathcal{A}_n^\circ,\, n \in \Delta)$ is tight (in distribution) in $\mathcal{M}_{\mathfrak{F}}(\mathbb{T} \times \mathbb{R}_+)$. Now apply Proposition A.10 and Proposition A.9 to get that

$$\mathcal{A}_n^\circ(\mathrm{f}h) \xrightarrow[n \to \infty]{(d)} \Psi_{\mathcal{T}}(\mathrm{f}h)$$

for every $h \in \mathcal{C}_b(\mathbb{T} \times \mathbb{R}_+)$ and every $\mathrm{f} \in \mathfrak{F}$. Let $f \in \mathcal{B}(\mathbb{T} \times \mathbb{R}_+)$ satisfying the assumptions of Proposition 7.1. Consider $\mathrm{f} = g_1$ and $h = f/g_1$. Notice that (7.1) implies that h is continuous on

$\mathbb{T} \times \mathbb{R}_+$. Since $\mathrm{f}h = g_1 h = f$ except possibly on $S_0 = \mathbb{T}_0 \times \mathbb{R}_+$ and $\mathcal{A}_n^\circ(S_0) = \Psi_\mathcal{T}(S_0) = 0$, we deduce that the convergence in distribution (7.2) holds.

Let $p > 1$ such that $p(\gamma\alpha + (\gamma-1)\beta) > 1-\gamma$. There exists $q > p$ satisfying the same inequality. Since $|f(T,r)| \leq C\mathfrak{m}(T)^\alpha \mathfrak{h}(T)^\beta$, we get that

$$\sup_{n\in\Delta} \mathbb{E}\left[|\mathcal{A}_n^\circ(f)|^q\right] \leq C^q \sup_{n\in\Delta} \mathbb{E}\left[\left(\frac{b_n^{1+\beta}}{n^{2+\alpha+\beta}} \sum_{w\in\tau^{n,\circ}} |\tau_w^n|^{1+\alpha}\mathfrak{h}(\tau_w^n)^\beta\right)^q\right], \tag{7.7}$$

where the right-hand side is finite by Lemma 5.2. Thus, the sequence $(|\mathcal{A}_n^\circ(f)|^p,\, n \in \Delta)$ is uniformly integrable and the convergence of the moment of order p of $\mathcal{A}_n^\circ(f)$ towards the moment of order p of $\Psi_\mathcal{T}(f)$ readily follows from (7.2). □

7.2. **Phase transition for functionals of the mass and height.** We refine the convergence result given in Proposition 7.1 for functionals depending only on the mass and height and describe a phase transition in that case.

We start with a technical lemma which is a consequence of the well-known de La Vallée Poussin criterion for uniform integrability.

Lemma 7.2. *Let ν be a nonnegative finite measure on $(0,1]$ and $f \in \mathcal{C}_+((0,1])$ be nonincreasing, belonging to $L^1(\nu)$ and such that $\lim_{x\to 0+} f(x) = +\infty$. Then there exists a positive function $f^\nu \in \mathcal{C}_+((0,1])$ which belongs to $L^1(\nu)$, such that f/f^ν is bounded on $(0,1]$ and $\lim_{x\to 0+} f(x)/f^\nu(x) = 0$.*

Proof. We may assume without loss of generality that f does not vanish anywhere in $(0,1]$ and that ν is a probability measure. By the de La Vallée Poussin criterion (see [13, §22]), there exists a convex nondecreasing function $F\colon \mathbb{R}_+ \to \mathbb{R}_+$ such that $\lim_{t\to\infty} F(t)/t = \infty$ and $F \circ f \in L^1(\nu)$. In fact, up to considering $F+1$ instead, we can and will assume that F does not vanish anywhere. Since F is convex on $\mathbb{R}_+$, it is continuous on $(0,\infty)$ and it follows that $F \circ f$ is continuous on $(0,1]$. Moreover, $F \circ f$ is clearly nonincreasing by composition. Further, since $\lim_{x\to 0} f(x) = \infty$ and $\lim_{t\to\infty} t/F(t) = 0$, we get $\lim_{x\to 0} f(x)/F\circ f(x) = 0$. The function $f/F \circ f$ being continuous on $(0,1]$ with a finite limit at 0, it is bounded on $(0,1]$. Setting $f^\nu = F \circ f$, the conclusion readily follows. □

We now give the main result of this section. Recall that the notation $\Psi_\mathcal{T}^{\mathfrak{mh}}(g(x)h(u))$ stands for $\Psi_\mathcal{T}^{\mathfrak{mh}}(f)$ where $f(x,u) = g(x)h(u)$. For $g \in \mathcal{B}(\mathbb{R}_+)$, define

$$g^*(x) \coloneqq \sup_{x\leq y\leq 1} |g(y)| \quad \text{for all } x \in (0,1]. \tag{7.8}$$

Theorem 7.3. *Assume that ξ satisfies (ξ1) and (ξ2)′.*

(i) Let $\beta \in \mathbb{R}$ and $g \in \mathcal{B}([0,1])$ be such that g is continuous on $(0,1]$ and satisfies

$$\int_0 g^*(x^{\gamma/(\gamma-1)})x^\beta \,\mathrm{d}x < \infty. \tag{7.9}$$

Then we have the convergence in distribution and of the first moment

$$\frac{b_n^{1+\beta}}{n^{2+\beta}} \sum_{w\in\tau^{n,\circ}} |\tau_w^n| \mathfrak{h}(\tau_w^n)^\beta g\left(\frac{|\tau_w^n|}{n}\right) \xrightarrow[n\to\infty]{(d)+\text{mean}} \Psi_\mathcal{T}^{\mathfrak{mh}}(g(x)u^\beta) \tag{7.10}$$

where $\Psi_{\mathcal{T}}^{\mathfrak{m}\mathfrak{h}}(|g(x)|u^{\beta})$ *is a.s. finite and integrable.*

(ii) Let $\alpha \in \mathbb{R}$ *and* $h \in \mathcal{B}(\mathbb{R}_+)$ *be such that* h *is continuous on* $(0, \infty)$ *and satisfies* $h(u) = O(\mathrm{e}^{u^{\eta}})$ *as* $u \to \infty$ *for some* $\eta \in (0, \gamma)$ *and*

$$\int_0 x^{\alpha\gamma/(\gamma-1)} h^*(x)\, \mathrm{d}x < \infty. \tag{7.11}$$

Then we have the convergence in distribution and of the first moment

$$\frac{b_n}{n^{2+\alpha}} \sum_{w \in \tau^{n,\circ}} |\tau_w^n|^{1+\alpha} h\left(\frac{b_n}{n}\mathfrak{h}(\tau_w^n)\right) \xrightarrow[n\to\infty]{(d)+\text{mean}} \Psi_{\mathcal{T}}^{\mathfrak{m}\mathfrak{h}}(x^{\alpha} h(u)) \tag{7.12}$$

where $\Psi_{\mathcal{T}}^{\mathfrak{m}\mathfrak{h}}(x^{\alpha}|h(u)|)$ *is a.s. finite and integrable.*

(iii) Let $f \in \mathcal{B}_+([0,1] \times \mathbb{R}_+)$ *be such that*

$$\int_0 f(x^{\gamma/(\gamma-1)}, x)\, \mathrm{d}x = \infty. \tag{7.13}$$

Suppose that f *is of the form* $f(x,u) = g(x)u^{\beta}$ *or* $f(x,u) = x^{\alpha}h(u)$ *where* $\alpha, \beta \in \mathbb{R}$ *and* g, h *are nonincreasing and continuous on* $(0,1]$ *and on* $(0, \infty)$ *respectively. Then we have*

$$\frac{b_n}{n^2} \sum_{w \in \tau^{n,\circ}} |\tau_w^n| f\left(\frac{|\tau_w^n|}{n}, \frac{b_n}{n}\mathfrak{h}(\tau_w^n)\right) \xrightarrow[n\to\infty]{(d)+\text{mean}} \infty. \tag{7.14}$$

Proof. To prove (i), we proceed in three steps.

Step 1 in the proof of (i). Let $g \in \mathcal{C}_+([0,1])$ be nonincreasing and nonzero. Let $(\beta_k,\, k \in \mathbb{N})$ be a decreasing sequence of nonpositive real numbers such that $\beta_0 = 0$ and $\lim_{k\to\infty} \beta_k = -1$. We define a set of functions $\mathfrak{F} = \{h_k \colon k \in \mathbb{N}\}$ where $h_k(u) = u^{\beta_k} \vee u^k$ for $u > 0$ and $k \in \mathbb{N}$, and $h_0(0) = 1$ and $h_k(0) = +\infty$ for $k \in \mathbb{N}$. We shall prove that $\mathfrak{F}$ satisfies assumptions (H1)–(H5) of Appendix A with $S = \mathbb{R}_+$ equipped with the Euclidean distance and $S_0 = \{0\}$. Notice that $h_0 \equiv 1$ and h_k is continuous on S_0^c for every $k \in \mathbb{N}$, so (H1) and (H2) are satisfied. Moreover, for every $k \in \mathbb{N}$, the function h_k/h_{k+1} is continuous on $(0, \infty)$ and we have

$$\lim_{u\to 0+} \frac{h_k(u)}{h_{k+1}(u)} = \lim_{u\to 0+} u^{\beta_k - \beta_{k+1}} = 0 \quad \text{and} \quad \lim_{u\to +\infty} \frac{h_k(u)}{h_{k+1}(u)} = \lim_{u\to +\infty} \frac{1}{u} = 0,$$

so that (H4) and (H5) are satisfied. Finally, since the set $\{x \in S \colon \rho(x, S_0) \geq \varepsilon,\, \rho(x,0) \leq M\} = [\varepsilon, M]$ is compact and h_k is continuous, it is bounded there and (H3) is satisfied. Define a (random) measure on $\mathbb{R}_+$ by setting

$$\zeta_n(h) = \frac{b_n}{n^2} \sum_{w \in \tau^{n,\circ}} |\tau_w^n| g\left(\frac{|\tau_w^n|}{n}\right) h\left(\frac{b_n}{n}\mathfrak{h}(\tau_w^n)\right) \tag{7.15}$$

for every $h \in \mathcal{B}_+(\mathbb{R}_+)$. By (4.23), ζ_n converges to ζ in distribution in $\mathcal{M}(\mathbb{R}_+)$ and $\mathbb{E}[\zeta_n(\bullet)]$ converges to $\mathbb{E}[\zeta(\bullet)]$ in $\mathcal{M}(\mathbb{R}_+)$ where ζ is defined by $\zeta(h) = \Psi_{\mathcal{T}}^{\mathfrak{m}\mathfrak{h}}(g(x)h(u))$. But, since we have $\int_0 g(x) x^{(\beta_k+1)(1-1/\gamma)-1}\, \mathrm{d}x < \infty$ for every $k \in \mathbb{N}$, Lemma 5.1-(i) gives

$$\sup_{n\in\Delta} \mathbb{E}[\zeta_n(h_k)] \leq \sup_{n\in\Delta} \mathbb{E}\left[\zeta_n(u^{\beta_k})\right] + \sup_{n\in\Delta} \mathbb{E}\left[\zeta_n(u^k)\right] < \infty \qquad \text{for all } k \in \mathbb{N}.$$

Thus, Corollary A.11 yields the convergence in distribution $\zeta_n \xrightarrow{(d)} \zeta$ in $\mathcal{M}_{\mathfrak{F}}$ as well as the convergence of the first moment $\mathbb{E}[\zeta_n(\bullet)] \to \mathbb{E}[\zeta(\bullet)]$ in $\mathcal{M}_{\mathfrak{F}}$. By Proposition A.9, this implies that for every $g \in \mathcal{C}_+([0,1])$ nonincreasing and every $\beta > -1$, we have

$$\frac{b_n^{1+\beta}}{n^{2+\beta}} \sum_{w \in \tau^{n,\circ}} |\tau_w^n| \mathfrak{h}(\tau_w^n)^\beta g\left(\frac{|\tau_w^n|}{n}\right) \xrightarrow[n\to\infty]{(d)+\text{mean}} \Psi_{\mathcal{T}}^{\mathfrak{mh}}(g(x)u^\beta). \tag{7.16}$$

Step 2 in the proof of (i). Now fix $\beta > -1$ and define the (random) measure ξ_n on $[0,1]$ by

$$\xi_n(g) = \frac{b_n^{1+\beta}}{n^{2+\beta}} \sum_{w \in \tau^{n,\circ}} |\tau_w^n| \mathfrak{h}(\tau_w^n)^\beta g\left(\frac{|\tau_w^n|}{n}\right), \tag{7.17}$$

for every $g \in \mathcal{B}_+([0,1])$. Notice that (7.16) can be rewritten as

$$\xi_n(g) \xrightarrow[n\to\infty]{(d)+\text{mean}} \xi(g) \tag{7.18}$$

for every $g \in \mathcal{C}_+([0,1])$ nonincreasing, where the measure ξ is defined by $\xi(g) = \Psi_{\mathcal{T}}^{\mathfrak{mh}}(g(x)u^\beta)$. Moreover, Lemma 5.1-(i) applied with $g \equiv 1$ gives $\sup_{n\in\Delta} \mathbb{E}[\xi_n(1)] < \infty$. As a consequence, by the Markov inequality, we have $\lim_{r\to\infty} \sup_{n\in\Delta} \mathbb{P}(\xi_n(1) > r) = 0$. Since $[0,1]$ is compact, this means that the sequence of random measures $(\xi_n,\, n \in \Delta)$ is tight in distribution in $\mathcal{M}([0,1])$, see [30, Theorem 4.10]. Hence, it is relatively compact by Prokhorov's theorem as the space $\mathcal{M}([0,1])$ is Polish for the weak topology. Let $\hat{\xi}$ be a limit point. Then we have $\xi(g) \overset{(d)}{=} \hat{\xi}(g)$ for every $g \in \mathcal{C}_+([0,1])$ nonincreasing. Therefore, we get that $\xi \overset{(d)}{=} \hat{\xi}$ and the sequence $(\xi_n,\, n \in \Delta)$ has only one limit point ξ. Since it is relatively compact, we deduce that ξ_n converges to ξ in distribution in $\mathcal{M}([0,1])$. A similar deterministic argument shows that $\mathbb{E}[\xi_n(\bullet)]$ converges to $\mathbb{E}[\xi(\bullet)]$ in $\mathcal{M}([0,1])$.

Step 3 in the proof of (i). Let $\beta > -1$ and $g \in \mathcal{B}([0,1])$ be continuous on $(0,1]$, nonzero and such that $\int_0 g^*(x) x^{(\beta+1)(1-1/\gamma)-1}\, \mathrm{d}x < \infty$. Set $g_0 \equiv 1$. If $\lim_{x\to 0} g^*(x) = \infty$, set $g_1 = g^* + 1$. If g^* has a finite limit at 0 (which is then positive), then there exists $\varepsilon > 0$ such that $\int_0 x^{-\varepsilon} g^*(x) x^{(\beta+1)(1-1/\gamma)-1}\, \mathrm{d}x < \infty$. We also have $\lim_{x\to 0+} x^{-\varepsilon} g^*(x) = \infty$ and the function $x \mapsto x^{-\varepsilon} g^*(x)$ is continuous and nonincreasing. In that case, we set $g_1(x) = x^{-\varepsilon} g^*(x) + 1$ for $x \in [0,1]$.

Define a set of functions $\mathfrak{F} = \{g_k\colon\ k \in \mathbb{N}\}$ as follows: for every $k \geq 1$, set $g_{k+1} = g_k^\nu$ which is given by Lemma 7.2 applied with the finite measure $\nu(\mathrm{d}x) = x^{(\beta+1)(1-1/\gamma)-1}\mathrm{d}x$. By construction, the sequence $\mathfrak{F}$ satisfies assumptions (H1)–(H4) of Appendix A with $S = [0,1]$, $S_0 = \{0\}$ and $\mathfrak{F}^\star(g_k) = \{g_j\colon\ j > k\}$ (notice (H3) is automatically satisfied as $[0,1]$ is compact). Notice that, by Lemma 7.2, for every $k \in \mathbb{N}$, the function g_k is continuous and nonincreasing on $(0,1]$ and satisfies $\int_0 g_k(x) x^{(\beta+1)(1-1/\gamma)-1}\, \mathrm{d}x < \infty$. So, by Lemma 5.1, we get that

$$\sup_{n\in\Delta} \mathbb{E}[\xi_n(g_k)] < \infty \quad \text{for all } k \in \mathbb{N}.$$

Now, Corollary A.11 applies and yields, in conjunction with Proposition A.9, the convergence in distribution and of the first moment

$$\xi_n(g_k \ell) \xrightarrow[n\to\infty]{(d)+\text{mean}} \xi(g_k \ell)$$

for every $k \in \mathbb{N}$ and $\ell \in \mathcal{C}([0,1])$. Now apply this with $k = 1$ and $\ell = g/g_1$. Notice that $g_1 \ell = g$ except possibly on $S_0 = \{0\}$. Since $\xi_n(S_0) = \xi(S_0) = 0$, we deduce that

$$\xi_n(g) \xrightarrow[n\to\infty]{(d)+\text{mean}} \xi(g).$$

This, together with Proposition 6.9, proves (i).

The proof of (ii) is quite similar so we only indicate the changes compared with (i).
Step 1 in the proof of (ii). Let $h \in \mathcal{C}_+(\mathbb{R}^+)$ be nonincreasing and nonzero.

Taking a decreasing sequence $(\alpha_k,\, k \in \mathbb{N})$ of nonpositive real numbers such that $\alpha_0 = 0$ and $\lim_{k\to\infty} \alpha_k = -1 + 1/\gamma$ and defining a set of functions $\mathfrak{F} = \{g_k\colon\ k \in \mathbb{N}\}$ by $g_k(x) = x^{\alpha_k}$, we can show that for every $h \in \mathcal{C}_+(\mathbb{R}_+)$ nonincreasing and every $\alpha > -1 + 1/\gamma$, we have

$$\frac{b_n}{n^{2+\alpha}} \sum_{w\in\tau^{n,\circ}} |\tau_w^n|^{1+\alpha} h\left(\frac{b_n}{n}\mathfrak{h}(\tau_w^n)\right) \xrightarrow[n\to\infty]{(d)+\text{mean}} \Psi_{\mathcal{T}}^{\mathfrak{m}\mathfrak{h}}(x^\alpha h(u)). \tag{7.19}$$

Step 2 in the proof of (ii). Fix $\alpha > -1 + 1/\gamma$ and define the (random) measure ξ_n on $\mathbb{R}_+$ by

$$\xi_n(h) = \frac{b_n}{n^{2+\alpha}} \sum_{w\in\tau^{n,\circ}} |\tau_w^n|^{1+\alpha} h\left(\frac{b_n}{n}\mathfrak{h}(\tau_w^n)\right), \tag{7.20}$$

for every $h \in \mathcal{B}_+(\mathbb{R}_+)$. Notice that (7.19) can be rewritten as

$$\xi_n(h) \xrightarrow[n\to\infty]{(d)+\text{mean}} \xi(h) \tag{7.21}$$

for every $h \in \mathcal{C}_+(\mathbb{R}_+)$ nonincreasing, where the measure ξ is defined by $\xi(h) = \Psi_{\mathcal{T}}^{\mathfrak{m}\mathfrak{h}}(x^\alpha h(u))$. Moreover, Lemma 5.1-(ii) applied with $h \equiv 1$ gives $\sup_{n\in\Delta} \mathbb{E}\left[\xi_n(1)\right] < \infty$. As a consequence, by the Markov inequality, we have $\lim_{r\to\infty} \sup_{n\in\Delta} \mathbb{P}\left(\xi_n(1) > r\right) = 0$. Fix $\beta > 0$ and let $r > 0$. Then, using the inequality $\mathbf{1}_{[r,\infty)}(u) \le (u/r)^\beta$ for every $u \ge 0$, we get

$$\sup_{n\in\Delta} \mathbb{E}\left[\xi_n([r,\infty))\right] \le \frac{1}{r^\beta} \sup_{n\in\Delta} \mathbb{E}\left[\mathcal{A}_n^{\mathfrak{m}\mathfrak{h},\circ}(x^\alpha u^\beta)\right].$$

Notice that the right-hand side is finite by Lemma 5.2 since $\gamma\alpha + (\gamma-1)(\beta+1) > 0$. We deduce that

$$\inf_{K\subset\mathbb{R}_+} \sup_{n\in\Delta} \mathbb{E}\left[\xi_n(K^c)\right] = 0,$$

where the infimum is taken over all compact subsets $K \subset \mathbb{R}_+$. By [30, Theorem 4.10], this means that the sequence of random measures $(\xi_n,\, n \in \Delta)$ is tight in distribution in $\mathcal{M}(\mathbb{R}_+)$. Following the end of step 2 for property (i), we are then able to show that ξ_n converges to ξ in distribution in $\mathcal{M}([0,\infty))$ and $\mathbb{E}\left[\xi_n(\bullet)\right]$ converges to $\mathbb{E}\left[\xi(\bullet)\right]$ in $\mathcal{M}([0,\infty))$.

Step 3 in the proof of (ii). Let $h \in \mathcal{B}(\mathbb{R}_+)$ be continuous on $(0,\infty)$ such that h^* is non-zero, $\int_0 h^*(u) u^{\alpha\gamma/(\gamma-1)}\, \mathrm{d}u < \infty$ and $h(u) = O(\mathrm{e}^{u^\eta})$ as $u \to \infty$ for some $\eta \in (0,\gamma)$. Set $h_0 \equiv 1$ and define a positive function $h_1 \in \mathcal{B}_+((0,\infty))$ in the following way. If $\lim_{u\to 0} h^*(u) = \infty$, set $h_1 = h^* + 1$ on $(0,1]$. If h^* has a finite limit at 0 (which is positive as h^* is non-zero), then $\alpha > -1 + 1/\gamma$, and thus there exists $\varepsilon > 0$ such that $\int_0 u^{-\varepsilon} h^*(u) u^{\alpha\gamma/(\gamma-1)}\, \mathrm{d}u < \infty$. Moreover, we have $\lim_{u\to 0} u^{-\varepsilon} h^*(u) = \infty$ and the function $u \mapsto u^{-\varepsilon} h^*(u)$ is continuous and nonincreasing. In that case, we set $h_1(u) = u^{-\varepsilon} h^*(u) + 1$ for $u \in (0,1]$. Now extend h_1 to a continuous function on $(0,\infty)$ such that $h_1(u) = \exp(u^{\eta_1})$ for $u \ge 2$ for some $\eta_1 \in (\eta,\gamma)$. Define a set of functions $\mathfrak{F} = \{h_k\colon\ k \in \mathbb{N}\}$ as follows. Let $(\eta_k,\, k \ge 2)$ be an increasing sequence in (η_1,γ). Recall that $\alpha > -1 + 1/\gamma$ so that the measure $\nu(\mathrm{d}u) = \mathbf{1}_{(0,1]}(u)\, u^{\alpha\gamma/(\gamma-1)}\mathrm{d}u$ is finite. For every $k \ge 1$, define $h_{k+1} \in \mathcal{B}_+([0,\infty))$ continuous and positive on $(0,\infty)$ and such that $h_{k+1} = h_k^\nu$ on $(0,1]$, with h_k^ν defined in Lemma 7.2, and $h_{k+1}(u) = \exp(u^{\eta_{k+1}})$ for $u \ge 2$. In particular, we have $\lim_{x\to 0+} h_k(x)/h_{k+1}(x) = \lim_{x\to+\infty} h_k(x)/h_{k+1}(x) = 0$. Then, it is easy to check that the sequence $\mathfrak{F}$ satisfies assumptions (H1)–(H5) of Appendix A with $S = \mathbb{R}_+$, $S_0 = \{0\}$ and $\mathfrak{F}^\star(h_k) = \{h_j\colon\ j > k\}$ for $k \in \mathbb{N}$. Notice that, by Lemma 7.2, for every $k \in \mathbb{N}$, the function h_k is continuous and nonincreasing on $(0,1]$ and satisfies $\int_0 h_k(u)u^{\alpha\gamma/(\gamma-1)}\,\mathrm{d}u < \infty$. So, by Lemma 5.1 (i) and (ii), we get that for all $k \in \mathbb{N}$, there exists a finite constant $C_k > 0$ such that

$$\sup_{n\in\Delta} \mathbb{E}\left[\xi_n(h_k)\right] \le \sup_{n\in\Delta} \mathbb{E}\left[\xi_n(h_k \mathbf{1}_{(0,1]})\right] + C_k \sup_{n\in\Delta} \mathbb{E}\left[\xi_n(\exp(u^{\eta_k})\mathbf{1}_{\{u\ge 1\}})\right] < \infty.$$

Now, Corollary A.11 applies and yields, in conjunction with Proposition A.9, the convergence in distribution and of the first moment

$$\xi_n(h_k f) \xrightarrow[n\to\infty]{(d)+\text{mean}} \xi(h_k f)$$

for every $k \in \mathbb{N}$ and every $f \in \mathcal{C}(\mathbb{R}_+)$. Taking $k = 1$ and $f = h/h_1$ proves (7.12) as $\xi_n(S_0) = \xi(S_0) = 0$. This, together with Proposition 6.9, proves (ii).

To prove (iii), notice that by (4.23) we have the convergence in distribution $\mathcal{A}_n^{\mathfrak{mh},\circ} \xrightarrow{(d)} \Psi_{\mathcal{T}}^{\mathfrak{mh}}$ in the space $\mathcal{M}([0,1]\times\mathbb{R}_+)$. Thanks to Skorokhod's representation theorem, we may assume that we have a.s. convergence. Thus, we get that a.s. for every $k \in \mathbb{N}$,

$$\lim_{n\to\infty} \frac{b_n}{n^2} \sum_{w\in\tau^{n,\circ}} |\tau_w^n| \left(f\left(\frac{|\tau_w^n|}{n}, \frac{b_n}{n}\mathfrak{h}(\tau_w^n) \right) \wedge k \right) = \Psi_{\mathcal{T}}^{\mathfrak{mh}}(f\wedge k).$$

Therefore, we have for $k \in \mathbb{N}$

$$\liminf_{n\to\infty} \frac{b_n}{n^2} \sum_{w\in\tau^{n,\circ}} |\tau_w^n| f\left(\frac{|\tau_w^n|}{n}, \frac{b_n}{n}\mathfrak{h}(\tau_w^n) \right) \ge \Psi_{\mathcal{T}}^{\mathfrak{mh}}(f\wedge k). \tag{7.22}$$

But by the monotone convergence theorem and Proposition 6.9, we have that a.s. $\lim_{k\to\infty} \Psi_{\mathcal{T}}^{\mathfrak{mh}}(f\wedge k) = \Psi_{\mathcal{T}}^{\mathfrak{mh}}(f) = \infty$. Thus, (7.14) follows from (7.22) by letting k go to infinity. □

Recall from (4.16) that we excluded the leaves to be able to consider functions taking infinite values on trees whose height vanishes. In the particular case where the function only blows up as the mass goes to zero, one can get rid of this restriction.

Remark 7.4. Recall the definition of the random measure $\mathcal{A}_n^{\mathfrak{mh},\circ} \in \mathcal{M}([0,1]\times\mathbb{R}_+)$:

$$\mathcal{A}_n^{\mathfrak{mh},\circ}(f) = \frac{b_n}{n^2} \sum_{w\in\tau^{n,\circ}} |\tau_w^n| f\left(\frac{|\tau_w^n|}{n}, \frac{b_n}{n}\mathfrak{h}(\tau_w^n) \right).$$

Similarly to the measure $\mathcal{A}_n^{\mathfrak{mh},\circ}$, we define the measure $\mathcal{A}_n^{\mathfrak{mh}} \in \mathcal{M}([0,1]\times\mathbb{R}_+)$, where the sum is over all the vertices (the internal vertices and the leaves): for $f \in \mathcal{B}_+([0,1]\times\mathbb{R}_+)$

$$\mathcal{A}_n^{\mathfrak{mh}}(f) = \frac{b_n}{n^2} \sum_{w\in\tau^{n}} |\tau_w^n| f\left(\frac{|\tau_w^n|}{n}, \frac{b_n}{n}\mathfrak{h}(\tau_w^n) \right).$$

Let $\beta \ge 0$ and $g \in \mathcal{B}([0,1])$ such that g is continuous on $(0,1]$ and $\int_0 g^*(x^{\gamma/(\gamma-1)})x^\beta\,\mathrm{d}x < \infty$. By Theorem 7.3-(i), we have

$$\mathcal{A}_n^{\mathfrak{mh},\circ}(g(x)u^\beta) \xrightarrow[n\to\infty]{(d)+\text{mean}} \Psi_{\mathcal{T}}^{\mathfrak{mh}}(g(x)u^\beta). \tag{7.23}$$

Now note that

$$\mathcal{A}_n^{\mathfrak{mh}}(g(x)u^\beta) = \frac{b_n^{1+\beta}}{n^{2+\beta}} \sum_{w\in\tau^n} |\tau_w^n| \mathfrak{h}(\tau_w^n)^\beta g\left(\frac{|\tau_w^n|}{n} \right)$$

makes sense when the function g blows up at 0. If $\beta > 0$, we have $\mathcal{A}_n^{\mathfrak{mh}}(g(x)u^\beta) = \mathcal{A}_n^{\mathfrak{mh},\circ}(g(x)u^\beta)$ since $\mathfrak{h}(\tau_w^n) = 0$ for every leaf $w \in \mathrm{Lf}(\tau^n)$. Thus we only need to consider the case $\beta = 0$. Then, using (4.2) and the fact that $|\mathrm{Lf}(\tau^n)| \leq n$ and that $|\tau_w^n| = 1$ for every $w \in \mathrm{Lf}(\tau^n)$, we have

$$\left|\mathcal{A}_n^{\mathfrak{mh}}(g(x)) - \mathcal{A}_n^{\mathfrak{mh},\circ}(g(x))\right| = \frac{b_n}{n^2}\left|\sum_{w\in\mathrm{Lf}(\tau^n)} |\tau_w^n| g\left(\frac{|\tau_w^n|}{n}\right)\right| \leq \overline{b} n^{-1+1/\gamma} g^*\left(\frac{1}{n}\right).$$

Since g^* is nonincreasing and satisfies $\int_0 g^*(x^{\gamma/(\gamma-1)})\,\mathrm{d}x < \infty$, it is straightforward to check that $g^*(x) = o(x^{1/\gamma-1})$ as $x \to 0$. Thus, we deduce that $\lim_{n\to\infty} \mathcal{A}_n^{\mathfrak{mh}}(g(x)u^\beta) - \mathcal{A}_n^{\mathfrak{mh},\circ}(g(x)u^\beta) = 0$ a.s. and in $L^1(\mathbb{P})$. As a consequence, the convergence (7.23) still holds if we replace $\mathcal{A}_n^{\mathfrak{mh},\circ}(g(x)u^\beta)$ by $\mathcal{A}_n^{\mathfrak{mh}}(g(x)u^\beta)$.

Similarly, let $\alpha > -1+1/\gamma$ and $h \in \mathcal{C}(\mathbb{R}_+)$ such that $h(u) = O(\mathrm{e}^{u^\eta})$ as $u \to \infty$ for some $\eta \in (0,\gamma)$. Then h^* is bounded near 0 and necessarily $\int_0 x^{\alpha\gamma/(\gamma-1)} h^*(x)\,\mathrm{d}x < \infty$. Thus, by Theorem 7.3, we have

$$\mathcal{A}_n^{\mathfrak{mh},\circ}(x^\alpha h(u)) \xrightarrow[n\to\infty]{(d)+\text{mean}} \Psi_{\mathcal{T}}^{\mathfrak{mh}}(x^\alpha h(u)). \tag{7.24}$$

Furthermore, using (4.2) we have

$$\left|\mathcal{A}_n^{\mathfrak{mh}}(x^\alpha h(u)) - \mathcal{A}_n^{\mathfrak{mh},\circ}(x^\alpha h(u))\right| = \frac{b_n}{n^{2+\alpha}}\,|\mathrm{Lf}(\tau^n)|\,|h(0)| \leq \overline{b} n^{-\alpha-1+1/\gamma}|h(0)|.$$

Thus, we deduce that $\lim_{n\to\infty} \mathcal{A}_n^{\mathfrak{mh}}(x^\alpha h(u)) - \mathcal{A}_n^{\mathfrak{mh},\circ}(x^\alpha h(u)) = 0$ a.s. and in $L^1(\mathbb{P})$ and the convergence (7.24) holds for $\mathcal{A}_n^{\mathfrak{mh}}(x^\alpha h(u))$.

Example 7.5. Fix $\alpha > -1+1/\gamma$ and set $g(x) = |\log(x)|x^\alpha$. It is clear that $\int_0 g(x^{\gamma/(\gamma-1)})\,\mathrm{d}x < \infty$, so by Theorem 7.3 we have the convergence in distribution

$$\mathcal{A}_n^{\mathfrak{mh},\circ}(g(x)) \xrightarrow[n\to\infty]{(d)} \Psi_{\mathcal{T}}^{\mathfrak{mh}}(g(x)).$$

But notice that

$$\begin{aligned}\mathcal{A}_n^{\mathfrak{mh},\circ}(g(x)) &= \frac{b_n \log(n)}{n^{2+\alpha}} \sum_{w\in\tau^{n,\circ}} |\tau_w^n|^{1+\alpha} - \frac{b_n}{n^{2+\alpha}} \sum_{w\in\tau^{n,\circ}} |\tau_w^n|^{1+\alpha} \log|\tau_w^n| \\ &= \log(n)\, \mathcal{A}_n^{\mathfrak{mh},\circ}(x^\alpha) - \frac{b_n}{n^{2+\alpha}} \sum_{w\in\tau^{n,\circ}} |\tau_w^n|^{1+\alpha} \log|\tau_w^n|\,.\end{aligned}$$

Again Theorem 7.3 gives the convergence in distribution $\mathcal{A}_n^{\mathfrak{mh},\circ}(x^\alpha) \xrightarrow{(d)} \Psi_{\mathcal{T}}^{\mathfrak{mh}}(x^\alpha)$. Therefore, we get the following asymptotic expansion in distribution

$$\frac{b_n}{n^{2+\alpha}} \sum_{w\in\tau^{n,\circ}} |\tau_w^n|^{1+\alpha} \log|\tau_w^n| \overset{(d)}{=} \log(n)\, \Psi_{\mathcal{T}}^{\mathfrak{mh}}(x^\alpha) - \Psi_{\mathcal{T}}^{\mathfrak{mh}}(|\log(x)|x^\alpha) + o(1).$$

Furthermore, since

$$\lim_{n\to\infty} \mathbb{E}\left[\mathcal{A}_n^{\mathfrak{mh},\circ}(g(x))\right] = \mathbb{E}\left[\Psi_{\mathcal{T}}^{\mathfrak{mh}}(g(x))\right] \quad \text{and} \quad \lim_{n\to\infty} \mathbb{E}\left[\mathcal{A}_n^{\mathfrak{mh},\circ}(x^\alpha)\right] = \mathbb{E}\left[\Psi_{\mathcal{T}}^{\mathfrak{mh}}(x^\alpha)\right],$$

we get the corresponding asymptotic expansion for the first moment

$$\frac{b_n}{n^{2+\alpha}}\, \mathbb{E}\left[\sum_{w\in\tau^{n,\circ}} |\tau_w^n|^{1+\alpha} \log|\tau_w^n|\right] = \log(n)\, \mathbb{E}\left[\Psi_{\mathcal{T}}^{\mathfrak{mh}}(x^\alpha)\right] - \mathbb{E}\left[\Psi_{\mathcal{T}}^{\mathfrak{mh}}(|\log(x)|x^\alpha)\right] + o(1).$$

Appendix A. A space of measures

Let (S, ρ) be a Polish metric space, $S_0 \subset S$ be a closed set in S and $0 \in S_0$ be a distinguished point. Denote by $\mathcal{K}$ the class of compact sets $K \subset S$. For any $x \in S$ and $A \subset S$, the distance from x to A is defined by $\rho(x, A) = \inf\{\rho(x, y) \colon\ y \in A\}$. Let $\mathfrak{F}$ be a countable set of measurable $[0, +\infty]$-valued functions on S satisfying the following assumptions:

(H1) The constant function $\mathbf{1}$ belongs to $\mathfrak{F}$.
(H2) All $f \in \mathfrak{F}$ are continuous on S_0^c.
(H3) All $f \in \mathfrak{F}$ are bounded away from zero and infinity on $\{x \in S \colon\ \rho(x, S_0) \geq \varepsilon,\ \rho(x, 0) \leq M\}$ for every $0 < \varepsilon < M < +\infty$.
(H4) For all $f \in \mathfrak{F}$, the set $\mathfrak{F}^\star(f) \subset \mathfrak{F}$ of functions $f^\star \in \mathfrak{F}$ such that $f/f^\star$ is bounded on S_0^c and $\lim_{\rho(x,S_0)\to 0+} f(x)/f^\star(x) = 0$ is non-empty.

Note that assumption (H3) is automatically satisfied when S is compact and every $f \in \mathfrak{F}$ is positive on S_0^c. Notice that (H4) implies that $\mathfrak{F}^\star(f)$ is infinitely countable for any $f \in \mathfrak{F}$. We shall write $f^\star$ for any element of $\mathfrak{F}^\star(f)$. By (H1) and (H4), we have $\lim_{\rho(x,S_0)\to 0+} \mathbf{1}^\star(x) = +\infty$. By convention, we take $\mathbf{1}^\star \equiv +\infty$ on S_0 and $f/f^\star \equiv 0$ on S_0 for every $f \in \mathfrak{F}$. We will occasionally need the following additional assumption:

(H5) S is compact or $\inf_{K\in\mathcal{K}} \sup_{x\in K^c} f(x)/f^\star(x) = 0$ for every $f \in \mathfrak{F}$ (and some $f^\star \in \mathfrak{F}^\star(f)$).

Denote by $\mathcal{M} = \mathcal{M}(S)$ the space of nonnegative finite measures on S endowed with the weak topology. Recall that $(\mathcal{M}, d_{\mathrm{BL}})$, with d_{BL} the bounded Lipschitz distance is a Polish metric space. If $\mu \in \mathcal{M}$ and $f \in \mathcal{B}_+(S)$, we write $f\mu$ for the measure $f(x)\mu(\mathrm{d}x)$. Set

$$\mathcal{M}_{\mathfrak{F}} = \mathcal{M}_{\mathfrak{F}}(S) \coloneqq \{\mu \in \mathcal{M} \colon\ \mu(f) < \infty \ \text{ for all } f \in \mathfrak{F}\}\,. \tag{A.1}$$

For $\mu \in \mathcal{M}_{\mathfrak{F}}$, we have $\mu(S_0) = 0$ (as $\mathbf{1}^\star \equiv +\infty$ on S_0) and $f\mu \in \mathcal{M}$ for every $f \in \mathfrak{F}$. In particular, since $(f/f^\star)f^\star = f$ on S_0^c, we have $(f/f^\star)f^\star\mu = f\mu$ for every $f \in \mathfrak{F}$ (and $f^\star \in \mathfrak{F}^\star(f)$). We say a sequence $(\mu_n,\ n \in \mathbb{N})$ of elements of $\mathcal{M}_{\mathfrak{F}}$ converges to $\mu \in \mathcal{M}_{\mathfrak{F}}$ if and only if $(f\mu_n,\ n \in \mathbb{N})$ converges to $f\mu$ in $\mathcal{M}$ for every $f \in \mathfrak{F}$. We consider the following distance $d_{\mathfrak{F}}$ on $\mathcal{M}_{\mathfrak{F}}$ which defines the same topology:

$$d_{\mathfrak{F}}(\mu, \nu) = \sum_{k\in\mathbb{N}} \frac{1}{2^k} \left(1 \wedge d_{\mathrm{BL}}\left(f_k\mu, f_k\nu\right)\right) \quad \text{for} \quad \mu, \nu \in \mathcal{M}_{\mathfrak{F}}, \tag{A.2}$$

where $\{f_k \colon\ k \in \mathbb{N}\}$ is an enumeration of $\mathfrak{F}$. (The choice of the enumeration is unimportant, as the corresponding distances all define the same topology on $\mathcal{M}_{\mathfrak{F}}$.) Notice that the mapping $\mu \mapsto f\mu$ is continuous from $\mathcal{M}_{\mathfrak{F}}$ to $\mathcal{M}$. In particular, taking $f = \mathbf{1}$ gives that every sequence which converges in $\mathcal{M}_{\mathfrak{F}}$ also converges in $\mathcal{M}$ to the same limit.

We shall see that the space $(\mathcal{M}_{\mathfrak{F}}, d_{\mathfrak{F}})$ is complete and separable (Proposition A.1) and give a complete description of its compact subsets (Proposition A.2). The main goal of this section is to give conditions which allow to strengthen a convergence in $\mathcal{M}$ to a convergence in $\mathcal{M}_{\mathfrak{F}}$ for deterministic measures (Corollary A.3) and then to extend this result to random measures (Proposition A.10 and Corollary A.11).

Proposition A.1. *The space $(\mathcal{M}_{\mathfrak{F}}, d_{\mathfrak{F}})$ is complete and separable.*

Proof. Let $(\mu_n,\, n \in \mathbb{N})$ be a Cauchy sequence in $\mathcal{M}_{\mathfrak{F}}$. Then, by definition of $d_{\mathfrak{F}}$, the sequence $(f\mu_n,\, n \in \mathbb{N})$ is Cauchy in $\mathcal{M}$ for every $f \in \mathfrak{F}$. By completeness of $\mathcal{M}$, for every $f \in \mathfrak{F}$, there exists a measure $\nu_f \in \mathcal{M}$ such that $\lim_{n\to\infty} f\mu_n = \nu_f$ in $\mathcal{M}$. We claim that $\nu_f(S_0) = 0$ for every $f \in \mathfrak{F}$. Indeed, fix $f \in \mathfrak{F}$ and $f^\star \in \mathfrak{F}^\star(f)$. As $f^\star \in \mathfrak{F}$, we have $\lim_{n\to\infty} f^\star\mu_n = \nu_{f^\star}$ in $\mathcal{M}$. By (H4), the function $f/f^\star$ is continuous and bounded on S, so that the mapping $\pi \mapsto (f/f^\star)\pi$ is continuous on $\mathcal{M}$. In particular, we have $\lim_{n\to\infty} f\mu_n = (f/f^\star)\nu_{f^\star}$ in $\mathcal{M}$. On the other hand, we have $\lim_{n\to\infty} f\mu_n = \nu_f$ in $\mathcal{M}$. We deduce that $\nu_f = (f/f^\star)\nu_{f^\star}$. It follows that $\nu_f(S_0) = 0$ since $f/f^\star = 0$ on S_0.

We set $\mu = \nu_{\mathbf{1}}$ so that $\lim_{n\to\infty} \mu_n = \mu$ in $\mathcal{M}$. Let $f \in \mathfrak{F}$. We shall prove that $f\mu = \nu_f$. Consider the closed set $F_k = \{f \geq 1/k\}$ for $k \in \mathbb{N}^*$. Notice that $F_k \subset \operatorname{int}(F_{k+1})$. Therefore, by Urysohn's lemma, there exists, for $k \in \mathbb{N}^*$, a continuous function $\chi_k \colon S \to [0,1]$ such that $\chi_k = 1$ on F_k and $\operatorname{supp}(\chi_k) \subset \operatorname{int}(F_{k+1})$. Notice that $(\chi_k f/f)\mu_n = \chi_k\mu_n$ since $(f/f) = 1$ on S_0^c and $\mu_n(S_0) = 0$. Since χ_k and χ_k/f are continuous and bounded, the mappings $\nu \mapsto \chi_k\nu$ and $\nu \mapsto (\chi_k/f)\nu$ are continuous from $\mathcal{M}$ to itself. We deduce that $\chi_k\mu = \lim_{n\to\infty} \chi_k\mu_n = \lim_{n\to\infty}(\chi_k/f)f\mu_n = (\chi_k/f)\nu_f$ in $\mathcal{M}$. Letting k go to infinity, as $\chi_k \uparrow \mathbf{1}$ on S_0^c since f is positive on S_0^c, and $\mu(S_0) = \nu_f(S_0) = 0$, we deduce (using the monotone convergence theorem) that $\mu = (1/f)\nu_f$ and thus $f\mu = \nu_f$. Since this holds for all $f \in \mathfrak{F}$, this proves that $\mu \in \mathcal{M}_{\mathfrak{F}}$ and that $\lim_{n\to\infty} f\mu_n = f\mu$ in $\mathcal{M}$ for every $f \in \mathfrak{F}$. Thus $\mathcal{M}_{\mathfrak{F}}$ is complete.

Next, define $F'_n = \{x \in S \colon\ \rho(x, S_0) \geq 1/n,\ \rho(x, 0) \leq n\}$. We will identify the space $\mathcal{M}(F'_n)$ with the subset of $\mathcal{M}$ consisting of the measures whose support lies in F'_n. Notice that F'_n is a Polish space (when endowed with the topology induced by ρ) as a closed subset of the Polish space S. In particular, the set $\mathcal{M}(F'_n)$ endowed with the bounded Lipschitz distance is a Polish space. Let $f \in \mathfrak{F}$. By (H3), the functions f and $1/f$ are both continuous and bounded on F'_n, so it is easy to check that the topology induced by $d_{\mathfrak{F}}$ on $\mathcal{M}(F'_n)$ coincides with the topology of weak convergence, *i.e.* the one induced by d_{BL}. Therefore, the space $(\mathcal{M}(F'_n), d_{\mathfrak{F}})$ is separable. To prove that $\mathcal{M}_{\mathfrak{F}}$ is separable, it suffices to show that $\mathcal{M}_{\mathfrak{F}}$ is equal to the completion of $\bigcup_{n\geq 1} \mathcal{M}(F'_n)$ with respect to d_{BL}. Notice that $F'_n \subset \operatorname{int}(F'_{n+1})$. Therefore, by Urysohn's lemma, there exists a continuous function $\chi'_n \colon S \to [0,1]$ such that $\chi'_n = 1$ on F'_n and $\operatorname{supp}(\chi'_n) \subset \operatorname{int}(F'_{n+1})$. Let $\mu \in \mathcal{M}_{\mathfrak{F}}$ and set $\mu_n = \chi'_n\mu$. Then it is clear that μ_n has support in F'_{n+1} and thus $\mu \in \mathcal{M}(F'_{n+1})$. Moreover, for every $f \in \mathfrak{F}$ and every nonnegative $h \in \mathcal{C}_b(S)$, we have

$$\mu_n(hf) = \mu(hf\chi'_n) \xrightarrow[n\to\infty]{} \mu(hf)$$

by the monotone convergence theorem, since $\chi'_n \uparrow \mathbf{1}_{S_0^c}$ and $\mu(S_0) = 0$. This proves that $(f\mu_n, n \in \mathbb{N})$ converges to $f\mu$ in $\mathcal{M}$ for every $f \in \mathfrak{F}$, thus $d_{\mathfrak{F}}(\mu_n, \mu) \to 0$. This concludes the proof. □

A set of measures $A \subset \mathcal{M}$ is said to be bounded if $\sup_{\mu\in A} \mu(\mathbf{1}) < \infty$. We now give a characterization of compactness in $\mathcal{M}_{\mathfrak{F}}$.

Proposition A.2. *Let $A \subset \mathcal{M}_{\mathfrak{F}}$.*

(i) A is relatively compact if and only if for every $f \in \mathfrak{F}$, the family $\{f\mu \colon\ \mu \in A\}$ of finite measures is bounded and tight.
(ii) If (H5) holds, then A is relatively compact if and only if for every $f \in \mathfrak{F}$, the family $\{f\mu \colon\ \mu \in A\}$ is bounded.

Proof. To prove (i), start by assuming that A is relatively compact. For every $\mu \in \mathcal{M}_{\mathfrak{F}}$ and every $f \in \mathfrak{F}$, set $F_f(\mu) = f\mu$. This defines a continuous mapping $F_f : \mathcal{M}_{\mathfrak{F}} \to \mathcal{M}$. It follows that the set

$$F_f(A) = \{f\mu : \mu \in A\}$$

is relatively compact in $\mathcal{M}$, *i.e.* it is bounded and tight by Prokhorov's theorem.

Conversely, let us assume that $\{f\mu : \mu \in A\}$ is bounded and tight in $\mathcal{M}$ for all $f \in \mathfrak{F}$. Let $(\mu_n, n \in \mathbb{N})$ be a sequence in A. Since the sequence of measures $(f\mu_n, n \in \mathbb{N})$ is bounded and tight, it is relatively compact in $\mathcal{M}$ for every $f \in \mathfrak{F}$. Therefore, by diagonal extraction, there exists a subsequence still denoted by $(f\mu_n, n \in \mathbb{N})$ which converges in $\mathcal{M}$ for every $f \in \mathfrak{F}$. By the same argument as in the proof of Proposition A.1, it follows that $(\mu_n, n \in \mathbb{N})$ converges in $\mathcal{M}_{\mathfrak{F}}$. This proves that A is relatively compact.

To prove (ii), assume that (H5) holds. The statement for a compact S follows immediately since a family of finite measures on a compact space is always tight. Now assume that S is not compact and let $A \subset \mathcal{M}_{\mathfrak{F}}$ such that the family $\{f\mu : \mu \in A\}$ is bounded for every $f \in \mathfrak{F}$. To prove that $A \subset \mathcal{M}_{\mathfrak{F}}$ is relatively compact, it is enough to show that $\{f\mu : \mu \in A\}$ is tight and to apply the first point. Let $f^\star \in \mathfrak{F}^\star(f)$, which appears in (H5), and $K \subset S$ be a compact subset. For every $\mu \in A$, since $\mu(S_0) = 0$, we have

$$\begin{aligned}
\int_{K^c} f(x)\,\mu(\mathrm{d}x) &= \int_{K^c} f(x)\mathbf{1}_{S_0^c}(x)\,\mu(\mathrm{d}x) \\
&= \int_{K^c} \frac{f(x)}{f^\star(x)}\mathbf{1}_{S_0^c}(x) f^\star(x)\,\mu(\mathrm{d}x) \\
&\leq \mu(f^\star)\,\sup_{K^c} \frac{f}{f^\star}\cdot
\end{aligned}$$

It follows that

$$\sup_{\mu\in A} \int_{K^c} f(x)\,\mu(\mathrm{d}x) \leq \sup_{\mu\in A} \mu(f^\star)\ \sup_{K^c} \frac{f}{f^\star},$$

and taking the infimum over all compact subsets $K \in \mathcal{K}$ yields, thanks to (H5)

$$\inf_{K\in\mathcal{K}} \sup_{\mu\in A} \int_{K^c} f(x)\,\mu(\mathrm{d}x) = 0,$$

i.e. the family $\{f\mu : \mu \in A\}$ is tight. This completes the proof. □

The next result gives sufficient conditions allowing to strengthen convergence in $\mathcal{M}$ to convergence in $\mathcal{M}_{\mathfrak{F}}$.

Corollary A.3. *Let $(\mu_n, n \in \mathbb{N})$ be a sequence of elements of $\mathcal{M}_{\mathfrak{F}}$ converging in $\mathcal{M}$ to some $\mu \in \mathcal{M}$. Then $\mu \in \mathcal{M}_{\mathfrak{F}}$ and $\lim_{n\to\infty} \mu_n = \mu$ in $\mathcal{M}_{\mathfrak{F}}$ under either of the following conditions:*

(i) $(f\mu_n, n \in \mathbb{N})$ is bounded and tight for every $f \in \mathfrak{F}$.
(ii) (H5) holds and $(f\mu_n, n \in \mathbb{N})$ is bounded for every $f \in \mathfrak{F}$.

Proof. Either condition guarantees that the sequence $(\mu_n, n \in \mathbb{N})$ is relatively compact in $\mathcal{M}_{\mathfrak{F}}$ by Proposition A.2. Let $\hat{\mu} \in \mathcal{M}_{\mathfrak{F}}$ be a limit point of $(\mu_n, n \in \mathbb{N})$. Then there exists a subsequence, still denoted by $(\mu_n, n \in \mathbb{N})$ such that $\lim_{n\to\infty} \mu_n = \hat{\mu}$ in $\mathcal{M}_{\mathfrak{F}}$. In particular, we have $\lim_{n\to\infty} \mu_n = \hat{\mu}$ in $\mathcal{M}$. Since $\lim_{n\to\infty} \mu_n = \mu$ in $\mathcal{M}$ by assumption, it follows that $\hat{\mu} = \mu$. This proves that $\mu \in \mathcal{M}_{\mathfrak{F}}$

and that $\lim_{n\to\infty} \mu_n = \mu$ in $\mathcal{M}_{\mathfrak{F}}$ since the sequence $(\mu_n,\, n \in \mathbb{N})$ is relatively compact in $\mathcal{M}_{\mathfrak{F}}$ and has only one limit point μ. □

The compactness criterion of Proposition A.2 yields a tightness criterion for random measures in $\mathcal{M}_{\mathfrak{F}}$.

Proposition A.4. *Let Ξ be a family of $\mathcal{M}_{\mathfrak{F}}$-valued random variables.*

(i) The family Ξ is tight (in distribution) in $\mathcal{M}_{\mathfrak{F}}$ if and only if for every $f \in \mathfrak{F}$, the family $\{f\xi\colon\, \xi \in \Xi\}$ is tight (in distribution) in $\mathcal{M}$, i.e. *if and only if*

$$\lim_{r\to\infty} \sup_{\xi\in\Xi} \mathbb{P}\left(\xi(f) > r\right) = 0 \tag{A.3}$$

and

$$\inf_{K\in\mathcal{K}} \sup_{\xi\in\Xi} \mathbb{E}\left[1 \wedge \int_{K^c} f(x)\xi(\mathrm{d}x)\right] = 0. \tag{A.4}$$

(ii) If (H5) holds, then Ξ is tight (in distribution) in $\mathcal{M}_{\mathfrak{F}}$ if and only if (A.3) holds for every $f \in \mathfrak{F}$.

Proof. To prove (i), assume that Ξ is tight in $\mathcal{M}_{\mathfrak{F}}$. Since the mapping $F_f\colon \mu \mapsto f\mu$ is continuous from $\mathcal{M}_{\mathfrak{F}}$ to $\mathcal{M}$ for every $f \in \mathfrak{F}$ and since tightness is preserved by continuous mappings, it follows that the family $F_f(\Xi) = \{f\xi\colon\, \xi \in \Xi\}$ is tight in $\mathcal{M}$ for every $f \in \mathfrak{F}$. The result now follows from Theorem 4.10 in [30].

Conversely, assume that (A.3) and (A.4) hold for all $f \in \mathfrak{F}$ and let $\varepsilon > 0$. Let $\{f_k\colon\, k \in \mathbb{N}^*\}$ be an enumeration of $\mathfrak{F}$. We set for $k \in \mathbb{N}^*$:

$$C_k = k\left(1 + \sup_{j\le k,\, f_k\in\mathfrak{F}^\star(f_j)} \|f_j/f_k\|_\infty\right),$$

with the convention that $\sup\emptyset = 0$. For every $k \in \mathbb{N}^*$, there exists $r_k > 0$ and a compact set $K_k \in \mathcal{K}$ such that

$$\sup_{\xi\in\Xi} \mathbb{P}\left(\xi(f_k) > r_k\right) \le \frac{\varepsilon}{2^k} \quad \text{and} \quad \sup_{\xi\in\Xi} \mathbb{E}\left[1 \wedge \int_{K_k^c} f_k(x)\xi(\mathrm{d}x)\right] \le \frac{\varepsilon}{C_k 2^k}\cdot$$

Set

$$A_\varepsilon = \bigcap_{k\in\mathbb{N}^*} \left\{\mu \in \mathcal{M}_{\mathfrak{F}}\colon\, \mu(f_k) \le r_k \text{ and } \int_{K_k^c} f_k(x)\mu(\mathrm{d}x) \le \frac{1}{C_k}\right\}.$$

Then for every $\xi \in \Xi$, we have

$$\begin{aligned}\mathbb{P}\left(\xi \in A_\varepsilon^c\right) &= \mathbb{P}\left(\exists k \in \mathbb{N}^*,\, \xi(f_k) > r_k \text{ or } \int_{K_k^c} f_k(x)\xi(\mathrm{d}x) > \frac{1}{C_k}\right)\\ &\le \sum_{k\in\mathbb{N}^*} \mathbb{P}\left(\xi(f_k) > r_k\right) + \sum_{k\in\mathbb{N}^*} \mathbb{P}\left(\int_{K_k^c} f_k(x)\xi(\mathrm{d}x) > \frac{1}{C_k}\right) \le 2\varepsilon,\end{aligned}$$

where in the last inequality we used that

$$\mathbb{P}\left(\int_{K_k^c} f(x)\xi(\mathrm{d}x) > \frac{1}{C_k}\right) = \mathbb{P}\left(1 \wedge \int_{K_k^c} f_k(x)\xi(\mathrm{d}x) > \frac{1}{C_k}\right) \le C_k\, \mathbb{E}\left[1 \wedge \int_{K_k^c} f_k(x)\xi(\mathrm{d}x)\right] \le \frac{\varepsilon}{2^k}\cdot$$

Thus, to prove that Ξ is tight in $\mathcal{M}_{\mathfrak{F}}$, it remains to show that $A_\varepsilon \subset \mathcal{M}_{\mathfrak{F}}$ is relatively compact. We have $\sup_{\mu\in A_\varepsilon} \mu(f_k) \leq r_k < \infty$ so that the family $\{f_k\mu\colon \mu \in A_\varepsilon\}$ is bounded for every $k \in \mathbb{N}^*$. Moreover, for every $i \geq k$ such that $f_i \in \mathfrak{F}^\star(f_k)$, we have

$$\sup_{\mu\in A_\varepsilon} \int_{K_i^c} f_k(x)\mu(\mathrm{d}x) \leq \|f_k/f_i\|_\infty \sup_{\mu\in A_\varepsilon} \int_{K_i^c} f_i(x)\mu(\mathrm{d}x) \leq \frac{1}{i}\cdot$$

This implies that $\inf_{K\in\mathcal{K}} \sup_{\mu\in A_\varepsilon} \int_{K^c} f_k(x)\mu(\mathrm{d}x) \leq 1/i$ for $i \geq k$ such that $f_i \in \mathfrak{F}^\star(f_k)$. Since there are infinitely many such i, we deduce that

$$\inf_{K\in\mathcal{K}} \sup_{\mu\in A_\varepsilon} \int_{K^c} f_k(x)\mu(\mathrm{d}x) = 0,$$

i.e. the family $\{f_k\mu\colon \mu \in A_\varepsilon\}$ is tight. As this holds for all $k \in \mathbb{N}^*$, we get by Proposition A.2 that A_ε is relatively compact in $\mathcal{M}_{\mathfrak{F}}$ (in fact, A_ε is compact as it is closed). This proves (i). The proof of (ii) is similar. □

We now give a sufficient condition for tightness in the space $\mathcal{M}_{\mathfrak{F}}$.

Corollary A.5. *Assume that (H5) holds. Let Ξ be a family of $\mathcal{M}_{\mathfrak{F}}$-valued random variables such that for every $f \in \mathfrak{F}$,*

$$\sup_{\xi\in\Xi} \mathbb{E}\left[\xi(f)\right] < \infty. \tag{A.5}$$

Then Ξ is tight (in distribution) in $\mathcal{M}_{\mathfrak{F}}$.

Proof. By the Markov inequality, we have for every $f \in \mathfrak{F}$,

$$\sup_{\xi\in\Xi} \mathbb{P}\left(\xi(f) > r\right) \leq \frac{1}{r} \sup_{\xi\in\Xi} \mathbb{E}\left[\xi(f)\right] \xrightarrow[r\to\infty]{} 0.$$

This proves that Ξ is tight in $\mathcal{M}_{\mathfrak{F}}$ by Proposition A.4-(ii). □

We denote by $\mathcal{B}$ (resp. $\mathcal{B}_{\mathfrak{F}}$) the Borel σ-field on $(\mathcal{M}, d_{\mathrm{BL}})$ (resp. on $(\mathcal{M}_{\mathfrak{F}}, d_{\mathfrak{F}})$). We also denote by $\mathcal{B}_{\mathrm{tr}} = \{A \cap \mathcal{M}_{\mathfrak{F}}\colon A \in \mathcal{B}\}$ the trace σ-field of $\mathcal{B}$ on $\mathcal{M}_{\mathfrak{F}}$.

Lemma A.6. *We have $\mathcal{B}_{\mathfrak{F}} = \mathcal{B}_{\mathrm{tr}}$.*

Proof. **Step 1.** We first prove that $\mathcal{M}_{\mathfrak{F}}$ is a Borel subset in $\mathcal{M}$. For $g \in \mathcal{B}_+(S)$, we consider the function Θ_g defined on $\mathcal{M}$ by $\Theta_g(\mu) = g\mu$. Denote $\mathcal{B}_{b+} = \mathcal{B}_b(S) \cap \mathcal{B}_+(S)$ the set of bounded nonnegative measurable functions defined on S. We follow the proof of [9, Theorem 15.13] to prove that, for every $g \in \mathcal{B}_{b+}$, Θ_g is a measurable function from $\mathcal{M}$ to $\mathcal{M}$. Denote by $\mathcal{F} = \{g \in \mathcal{B}_{b+}\colon \Theta_g \text{ is measurable}\}$. The function Θ_g is continuous for g belonging to $\mathcal{C}_{b+} = \mathcal{C}_b(S) \cap \mathcal{C}_+(S)$. Furthermore, the set $\mathcal{F}$ is closed under bounded pointwise convergence: if $g_n \to g$ pointwise, with $g \in \mathcal{B}_{b+}$ and $(g_n, n \in \mathbb{N})$ a bounded sequence of elements of $\mathcal{F}$ (*i.e.* $\sup_{n\in\mathbb{N}} \|g_n\|_\infty < \infty$), then $\Theta_g(\mu) = \lim_{n\to\infty} \Theta_{g_n}(\mu)$ by dominated convergence and thus g belongs to $\mathcal{F}$. An immediate extension of [9, Theorem 4.33] gives that $\mathcal{B}_{b+} \subset \mathcal{F}$.

We then deduce that the function $\theta_g\colon \mathcal{M} \to [0, +\infty]$ defined by $\theta_g(\mu) = g\mu(\mathbf{1}) = \mu(g)$ is measurable for every $g \in \mathcal{B}_{b+}$, and as $g \in \mathcal{B}_+(S)$ is the limit of $g \wedge n \in \mathcal{B}_{b+}$ as n goes to infinity, we deduce by monotone convergence that $\theta_g = \lim_{n\to\infty} \theta_{g\wedge n}$, and thus θ_g is measurable for every $g \in \mathcal{B}_+(S)$. By definition of $\mathcal{M}_{\mathfrak{F}}$, we have that $\mathcal{M}_{\mathfrak{F}} = \bigcap_{f\in\mathfrak{F}} \theta_f^{-1}(\mathbb{R}_+)$, and thus $\mathcal{M}_{\mathfrak{F}}$ is a Borel subset in $\mathcal{M}$.

Step 2. We prove that for every $\mu \in \mathcal{M}_{\mathfrak{F}}$, the mapping $\nu \mapsto d_{\mathfrak{F}}(\mu, \nu)$ defined on $\mathcal{M}_{\mathfrak{F}}$ is $\mathcal{B}_{\text{tr}}$-measurable. Let $g \in \mathcal{B}_{b+}$. Since the function Θ_g is measurable from $\mathcal{M}$ to itself by step 1, it is $\mathcal{B}/\mathcal{B}$-measurable. By definition of the trace σ-field, it follows that the mapping Θ_g from $\mathcal{M}_{\mathfrak{F}}$ to $\mathcal{M}$ is $\mathcal{B}_{\text{tr}}/\mathcal{B}$-measurable. Let $f \in \mathfrak{F}$. By monotone convergence we get that $\Theta_f = \lim_{n\to\infty} \Theta_{f\wedge n}$, and thus Θ_f is $\mathcal{B}_{\text{tr}}/\mathcal{B}$-measurable.

Since $\mu \in \mathcal{M}_{\mathfrak{F}}$, we have $f\mu \in \mathcal{M}$ and the mapping $\pi \mapsto d_{\text{BL}}(f\mu, \pi)$ from $\mathcal{M}$ to $\mathbb{R}$ is continuous hence $\mathcal{B}$-measurable. Thus, by composition we get that the mapping $\nu \mapsto d_{\text{BL}}(f\mu, f\nu)$ from $\mathcal{M}_{\mathfrak{F}}$ to $\mathbb{R}$ is $\mathcal{B}_{\text{tr}}$-measurable. Finally, the mapping $\nu \mapsto d_{\mathfrak{F}}(\mu, \nu)$ from $\mathcal{M}_{\mathfrak{F}}$ to $\mathbb{R}$ is $\mathcal{B}_{\text{tr}}$-measurable as a sum of $\mathcal{B}_{\text{tr}}$-measurable mappings.

Step 3. We conclude the proof of the lemma. For every $\mu \in \mathcal{M}_{\mathfrak{F}}$ and every $\varepsilon > 0$, we have

$$B(\mu, \varepsilon) = \{\nu \in \mathcal{M}_{\mathfrak{F}} \colon\ d_{\mathfrak{F}}(\mu, \nu) < \varepsilon\} \in \mathcal{B}_{\text{tr}}$$

by Step 2. Since $\mathcal{M}_{\mathfrak{F}}$ is a Polish space, every open set is the countable union of open balls and it follows that every open set lies in $\mathcal{B}_{\text{tr}}$. Hence we get $\mathcal{B}_{\mathfrak{F}} \subset \mathcal{B}_{\text{tr}}$.

Conversely, notice that the identity mapping from $(\mathcal{M}_{\mathfrak{F}}, d_{\mathfrak{F}})$ to $(\mathcal{M}_{\mathfrak{F}}, d_{\text{BL}})$ is continuous. Therefore, if $V \subset \mathcal{M}$ is an open set, $V \cap \mathcal{M}_{\mathfrak{F}}$ is open in $(\mathcal{M}_{\mathfrak{F}}, d_{\text{BL}})$ hence also in $(\mathcal{M}_{\mathfrak{F}}, d_{\mathfrak{F}})$. In particular, we have $V \cap \mathcal{M}_{\mathfrak{F}} \in \mathcal{B}_{\mathfrak{F}}$. Since this is true for every open set $V \subset \mathcal{M}$, we deduce that $\mathcal{B}_{\text{tr}} \subset \mathcal{B}_{\mathfrak{F}}$. ☐

The following two results are a direct consequence of Lemma A.6.

Corollary A.7. *Let ξ be a $\mathcal{M}$-valued random variable such that a.s. $\xi(f) < \infty$ for every $f \in \mathfrak{F}$. Then ξ is a $\mathcal{M}_{\mathfrak{F}}$-valued random variable. Conversely, if ξ is a $\mathcal{M}_{\mathfrak{F}}$-valued random variable then ξ is also a $\mathcal{M}$-valued random variable.*

Corollary A.8. *Let ξ and ζ be $\mathcal{M}_{\mathfrak{F}}$-valued random variables. Then the following conditions are equivalent:*

(i) $\xi \overset{(d)}{=} \zeta$ when viewed as $\mathcal{M}_{\mathfrak{F}}$-valued random variables.
(ii) $\xi \overset{(d)}{=} \zeta$ when viewed as $\mathcal{M}$-valued random variables.
(iii) $\xi(h) \overset{(d)}{=} \zeta(h)$ for every $h \in \mathcal{C}_b(S)$.
(iv) $\xi(fh) \overset{(d)}{=} \zeta(fh)$ for every $h \in \mathcal{C}_b(S)$ and $f \in \mathfrak{F}$.

We now characterize convergence in distribution of random measures in $\mathcal{M}_{\mathfrak{F}}$. Recall that (H1)–(H4) are in force.

Proposition A.9. *Let ξ_n and ξ be $\mathcal{M}_{\mathfrak{F}}$-valued random variables. Then ξ_n converges in distribution to ξ in $\mathcal{M}_{\mathfrak{F}}$ if and only if $\xi_n(fh) \xrightarrow[n\to\infty]{(d)} \xi(fh)$ for every $h \in \mathcal{C}_b(S)$ and every $f \in \mathfrak{F}$.*

Proof. Assume that ξ_n converges in distribution to ξ in $\mathcal{M}_{\mathfrak{F}}$. Let $f \in \mathfrak{F}$. Since $F \colon \mu \mapsto f\mu$ is continuous from $\mathcal{M}_{\mathfrak{F}}$ to $\mathcal{M}$ and $\nu \mapsto \nu(h)$ is continuous from $\mathcal{M}$ to $\mathbb{R}$ for every $h \in \mathcal{C}_b(S)$, it follows that the mapping $\mu \mapsto \mu(fh)$ is continuous from $\mathcal{M}_{\mathfrak{F}}$ to $\mathbb{R}$. By the continuous mapping theorem, we get $\xi_n(fh) \xrightarrow{(d)} \xi(fh)$.

Conversely, for every $f \in \mathfrak{F}$, $f\xi_n$ and $f\xi$ are $\mathcal{M}$-valued random variables, and we have $\xi_n(fh) \xrightarrow{(d)} \xi(fh)$ for every $h \in \mathcal{C}_b(S)$. By [30, Theorem 4.11], this implies that $f\xi_n \xrightarrow[n\to\infty]{(d)} f\xi$ in the space $\mathcal{M}$. In particular, $(f\xi_n,\, n \in \mathbb{N})$ is tight (in distribution) in $\mathcal{M}$ for every $f \in \mathfrak{F}$. By Proposition A.4, it follows that $(\xi_n,\, n \in \mathbb{N})$ is tight in $\mathcal{M}_{\mathfrak{F}}$. Since $\mathcal{M}_{\mathfrak{F}}$ is Polish, Prokhorov's theorem ensures that $(\xi_n,\, n \in \mathbb{N})$ is relatively compact (in distribution) in $\mathcal{M}_{\mathfrak{F}}$. Let $\hat{\xi}$ be a limit point (in distribution) of $(\xi_n,\, n \in \mathbb{N})$. There exists a subsequence, still denoted by ξ_n, such that $\xi_n \xrightarrow{(d)} \hat{\xi}$ in $\mathcal{M}_{\mathfrak{F}}$. Let $h \in \mathcal{C}_b(S)$. Applying the first part of the proof, we get that $\xi_n(fh) \xrightarrow[n\to\infty]{(d)} \hat{\xi}(fh)$ for every $f \in \mathfrak{F}$. Therefore, we have $\hat{\xi}(fh) \overset{(d)}{=} \xi(fh)$ for every $h \in \mathcal{C}_b(S)$. It follows from Corollary A.8 that $\hat{\xi} \overset{(d)}{=} \xi$ in $\mathcal{M}_{\mathfrak{F}}$. Thus the sequence $(\xi_n,\, n \in \mathbb{N})$ is relatively compact and has only one limit point ξ in $\mathcal{M}_{\mathfrak{F}}$. This proves the result. □

We state now the main result of this section. Recall that (H1)–(H4) are in force.

Proposition A.10. *Let $(\xi_n,\, n \in \mathbb{N})$ be a sequence of $\mathcal{M}_{\mathfrak{F}}$-valued random variables and ξ be a $\mathcal{M}$-valued random variable such that $\xi_n \xrightarrow{(d)} \xi$ in $\mathcal{M}$ and $(\xi_n, n \in \mathbb{N})$ is tight (in distribution) in $\mathcal{M}_{\mathfrak{F}}$. Then ξ is a $\mathcal{M}_{\mathfrak{F}}$-valued random variable and we have the convergence in distribution $\xi_n \xrightarrow{(d)} \xi$ in $\mathcal{M}_{\mathfrak{F}}$.*

Proof. By assumption, the sequence $(\xi_n,\, n \in \mathbb{N})$ is relatively compact (in distribution) in the space $\mathcal{M}_{\mathfrak{F}}$. Let $\hat{\xi} \in \mathcal{M}_{\mathfrak{F}}$ be a limit point in distribution and let $h \in \mathcal{C}_b(S)$. On the one hand, Proposition A.9 applied with $f = \mathbf{1}$ yields the convergence $\xi_n(h) \xrightarrow{(d)} \hat{\xi}(h)$. On the other hand, since $\xi_n \xrightarrow{(d)} \xi$ in $\mathcal{M}$ it follows that $\xi_n(h) \xrightarrow{(d)} \xi(h)$. Therefore $\hat{\xi}(h) \overset{(d)}{=} \xi(h)$ for every $h \in \mathcal{C}_b(S)$, *i.e.* $\hat{\xi} \overset{(d)}{=} \xi$ in $\mathcal{M}$. Since the distribution of $\hat{\xi}$ is concentrated on $\mathcal{M}_{\mathfrak{F}}$, the same is true for ξ. In other words $\xi \in \mathcal{M}_{\mathfrak{F}}$ a.s., and so ξ is a $\mathcal{M}_{\mathfrak{F}}$-valued random variable by Corollary A.7. Now, applying Corollary A.8 we get $\hat{\xi} \overset{(d)}{=} \xi$ in the space $\mathcal{M}_{\mathfrak{F}}$. Thus the sequence $(\xi_n,\, n \in \mathbb{N})$ is relatively compact in $\mathcal{M}_{\mathfrak{F}}$ and has only one limit point ξ, so $\xi_n \xrightarrow{(d)} \xi$ in $\mathcal{M}_{\mathfrak{F}}$. □

The following special case is particularly useful. Recall that (H1)–(H4) are in force.

Corollary A.11. *Assume that (H5) holds. Let $(\xi_n,\, n \in \mathbb{N})$ and ξ be $\mathcal{M}$-valued random variables such that $\xi_n \xrightarrow{(d)} \xi$ in $\mathcal{M}$ and for every $f \in \mathfrak{F}$,*

$$\sup_n \mathbb{E}\left[\xi_n(f)\right] < \infty. \tag{A.6}$$

Then $(\xi_n,\, n \in \mathbb{N})$ and ξ are $\mathcal{M}_{\mathfrak{F}}$-valued random variables and we have the convergence in distribution $\xi_n \xrightarrow{(d)} \xi$ in $\mathcal{M}_{\mathfrak{F}}$. Moreover, for every $f \in \mathfrak{F}$, we have

$$\mathbb{E}\left[\xi(f)\right] \leq \liminf_{n\to\infty} \mathbb{E}\left[\xi_n(f)\right] < \infty.$$

Furthermore, if $(\mathbb{E}\left[\xi_n(\bullet)\right],\, n \in \mathbb{N})$ converges to $\mathbb{E}\left[\xi(\bullet)\right]$ in $\mathcal{M}$ then the convergence actually holds in $\mathcal{M}_{\mathfrak{F}}$.

Proof. The random variable ξ_n is $\mathcal{M}$-valued and satisfies $\xi_n(f) < \infty$ a.s. since $\mathbb{E}\left[\xi_n(f)\right] < \infty$ for every $f \in \mathfrak{F}$, so by Corollary A.7, ξ_n is a $\mathcal{M}_{\mathfrak{F}}$-valued random variable. By Corollary A.5, the assumption (A.6) implies that $(\xi_n,\, n \in \mathbb{N})$ is tight (in distribution) in $\mathcal{M}_{\mathfrak{F}}$. Therefore Proposition

A.10 applies and gives the convergence in distribution $\xi_n \xrightarrow{(d)} \xi$ in $\mathcal{M}_{\mathfrak{F}}$. Moreover, Skorokhod's representation theorem in conjunction with Fatou's lemma implies that for every $f \in \mathfrak{F}$,

$$\mathbb{E}\left[\xi(f)\right] \leq \liminf_{n\to\infty} \mathbb{E}\left[\xi_n(f)\right] < \infty.$$

Now set $\mu_n = \mathbb{E}\left[\xi_n(\bullet)\right]$ and $\mu = \mathbb{E}\left[\xi(\bullet)\right]$ and assume that $\mu_n \to \mu$ in $\mathcal{M}$. Notice that the assumption (A.6) implies that $\mu_n \in \mathcal{M}_{\mathfrak{F}}$ for every $n \in \mathbb{N}$ and that the sequence of measures $(f\mu_n,\, n \in \mathbb{N})$ is bounded for every $f \in \mathfrak{F}$. Thus Corollary A.3 gives the convergence $\lim_{n\to\infty} \mu_n = \mu$ in $\mathcal{M}_{\mathfrak{F}}$. ☐

Appendix B. Sub-exponential tail bounds for the height of conditioned BGW trees

Assume that ξ satisfies $(\xi 1)$ and $(\xi 2)$ and denote by τ^n a BGW(ξ) tree conditioned to have n vertices. Then by [34, Theorem 1] which is stated for the aperiodic case but is trivially extended to the general case, for every $\alpha \in (0, \gamma/(\gamma-1))$, there exist two constants $C_0, c_0 > 0$ such that for every $y \geq 0$ and every $n \in \Delta$

$$\mathbb{P}\left(\frac{b_n}{n}\mathfrak{h}(\tau^n) \leq y\right) \leq C_0 \exp\left(-c_0 y^{-\alpha}\right). \tag{B.1}$$

We will show that under the stronger assumption $(\xi 2)'$, the previous inequality holds with $\alpha = \gamma/(\gamma-1)$. Since the finite variance case has already been treated in [5], we assume henceforth that ξ has infinite variance.

Recall that L is a slowly varying function such that $\mathbb{E}\left[\xi^2 \mathbf{1}_{\{\xi\leq n\}}\right] = n^{2-\gamma}L(n)$. On the other hand, the slowly varying function appearing in the appendix of [34], which we denote by K, satisfies $\mathrm{Var}\left(\xi \mathbf{1}_{\{\xi\leq n\}}\right) = n^{2-\gamma}K(n)$. Since $\mathrm{Var}(\xi) = +\infty$, we have as n goes to infinity that

$$\mathbb{E}\left[\xi^2 \mathbf{1}_{\{\xi\leq n\}}\right] \sim n^{2-\gamma}K(n) + 1 \sim n^{2-\gamma}K(n),$$

see the appendix in [34]. Therefore, we get $K(n) \sim L(n)$ and K is bounded above.

Following the proof of [34, Theorem 1] to get (B.1) holds for $\alpha = \gamma/(\gamma-1)$, it is enough to prove the analogue of Proposition 8 therein with $\alpha = \gamma/(\gamma-1)$, that is Proposition B.1 below. Let $(W_n,\, n \in \mathbb{N})$ be a random walk with starting point $W_0 = 0$ and jump distribution $\xi - 1$.

Proposition B.1. *Assume that ξ satisfies $(\xi 1)$ and $(\xi 2)'$. There exist two constants $C_0, c_0 > 0$ such that for every $u \geq 0$ and every $n \geq 1$,*

$$\mathbb{P}\left(\min_{1\leq i\leq n} W_i \leq -ub_n\right) \leq C_0 \exp\left(-c_0 u^{\gamma/(\gamma-1)}\right). \tag{B.2}$$

Proof. Note that $\mathbb{P}\left(\min_{1\leq i\leq n} W_i \leq -ub_n\right) = 0$ if $ub_n > n$, so that it is enough to prove (B.2) for $1 \leq u \leq n/b_n$. Write, for $h > 0$

$$\mathbb{P}\left(\min_{1\leq i\leq n} W_i \leq -ub_n\right) = \mathbb{P}\left(\max_{1\leq i\leq n} \mathrm{e}^{-hW_i} \geq \mathrm{e}^{hub_n}\right) \leq \mathrm{e}^{-hub_n}\,\mathbb{E}\left[\mathrm{e}^{-hW_n}\right] = \mathrm{e}^{-hub_n}\,\mathbb{E}\left[\mathrm{e}^{-hW_1}\right]^n, \tag{B.3}$$

where the inequality follows from Doob's maximal inequality applied to the submartingale $(\mathrm{e}^{-hW_n},\, n \in \mathbb{N})$. We shall apply (B.3) with $h = \varepsilon u^\eta / b_n$ where $\eta = 1/(\gamma-1)$ and $\varepsilon > 0$ is a constant to be chosen later. Note that $\gamma/(\gamma-1)=\eta\gamma=1+\eta$. Observe that $\varepsilon u^\eta/b_n$ is bounded uniformly in $1\leq u\leq n/b_n$ and $n\geq 1$. Indeed, since $b_n\geq \underline{b}n^{1/\gamma}$, we have

$$\frac{u^\eta}{b_n}\leq\left(\frac{n}{b_n}\right)^\eta\frac{1}{b_n}\leq\frac{1}{\underline{b}^{1+\eta}}.$$

Therefore, by [34, Eq. (42)], we have for every $n\geq 1$ and every $1\leq u\leq n/b_n$

$$\mathbb{E}\left[\mathrm{e}^{-\varepsilon\frac{u^\eta}{b_n}W_1}\right]\leq\exp\left\{Cn\left(\varepsilon\frac{u^\eta}{b_n}\right)^\gamma K\left(\frac{b_n}{\varepsilon u^\eta}\right)\right\}\leq\exp\left(C'\varepsilon^\gamma u^{\eta\gamma}\right),$$

as K is bounded from above and $b_n\geq\underline{b}n^{1/\gamma}$. Thus, we deduce from (B.3) that for $1\leq u\leq b_n/n$

$$\mathbb{P}\left(\min_{1\leq i\leq n}W_i\leq -ub_n\right)\leq\exp\left(-\left(\varepsilon-C'\varepsilon^\gamma\right)u^{1+\eta}\right).$$

The conclusion readily follows by choosing $\varepsilon>0$ small enough such that $\varepsilon-C'\varepsilon^\gamma>0$. $\square$

Remark B.2. In fact, this proof is valid if we only assume that the slowly varying function L of $(\xi 2)'$ is bounded from above, in which case $n^{-1/\gamma}b_n$ is bounded below.

References

[1] R. Abraham and J.-F. Delmas, *The forest associated with the record process on a Lévy tree*, Stochastic Processes and their Applications **123** (2013), no. 9, 3497–3517.

[2] R. Abraham, J.-F. Delmas, and P. Hoscheit, *A note on the Gromov-Hausdorff-Prokhorov distance between (locally) compact metric measure spaces*, Electronic Journal of Probability **18** (2013).

[3] R. Abraham and L. Serlet, *Poisson snake and fragmentation*, Electron. J. Probab. **7** (2002), 15 pp.

[4] L. Addario-Berry, N. Broutin, C. Goldschmidt, and G. Miermont, *The scaling limit of the minimum spanning tree of the complete graph*, The Annals of Probability **45** (2017), no. 5, 3075–3144.

[5] L. Addario-Berry, L. Devroye, and S. Janson, *Sub-Gaussian tail bounds for the width and height of conditioned Galton–Watson trees*, The Annals of Probability **41** (2013), no. 2, 1072–1087.

[6] P.-M. Agapow and A. Purvis, *Power of eight tree shape statistics to detect nonrandom diversification: a comparison by simulation of two models of cladogenesis*, Systematic biology **51** (2002), no. 6, 866–872.

[7] D. Aldous, *The continuum random tree. I*, The Annals of Probability (1991), 1–28.

[8] D. Aldous and J. Pitman, *The standard additive coalescent*, Ann. Probab. **26** (199810), no. 4, 1703–1726.

[9] C. D. Aliprantis and K. C. Border, *Infinite dimensional analysis*, Third, Springer, Berlin, 2006. A hitchhiker's guide.

[10] K. B. Athreya and P. E. Ney, *Branching processes*, Springer-Verlag, New York-Heidelberg, 1972. Die Grundlehren der mathematischen Wissenschaften, Band 196. MR0373040

[11] P. Biane, J. Pitman, and M. Yor, *Probability laws related to the Jacobi theta and Riemann zeta functions, and Brownian excursions*, Bulletin of the American Mathematical Society **38** (2001), no. 4, 435–465.

[12] V. I. Bogachev, *Measure theory*, Vol. 2, Springer Science & Business Media, 2007.

[13] C. Dellacherie and P.-A. Meyer, *Probabilités et potentiel: Chapitres i à iv*, Hermann, Paris, 1975.

[14] J.-F. Delmas, J.-S. Dhersin, and M. Sciauveau, *Cost functionals for large (uniform and simply generated) random trees*, Electron. J. Probab. **23** (2018), 36 pp.

[15] T. Duquesne, *A limit theorem for the contour process of condidtioned Galton-Watson trees*, The Annals of Probability **31** (2003), no. 2, 996–1027.

[16] T. Duquesne and J.-F. Le Gall, *Random trees, Lévy processes and spatial branching processes*, Astérisque **281** (2002), vi+147. MR1954248

[17] T. Duquesne and J.-F. Le Gall, *Probabilistic and fractal aspects of Lévy trees*, Probability Theory and Related Fields **131** (2005), no. 4, 553–603.

[18] T. Duquesne and M. Wang, *Decomposition of Lévy trees along their diameter*, Annales de l'Institut Henri Poincaré, Probabilités et Statistiques, 2017, pp. 539–593.

[19] K. B. Erickson and R. A. Maller, *Finiteness of integrals of functions of Lévy processes*, Proceedings of the London Mathematical Society **94** (2007), no. 2, 386–420.

[20] S. N. Evans, *Probability and real trees*, Lecture Notes in Mathematics, vol. 1920, Springer, Berlin, 2008. Lectures from the 35th Summer School on Probability Theory held in Saint-Flour, July 6–23, 2005. MR2351587

[21] P.-H. Fabre, L. Hautier, D. Dimitrov, and E. J. Douzery, *A glimpse on the pattern of rodent diversification: a phylogenetic approach*, BMC evolutionary biology **12** (2012), no. 1, 88.

[22] W. Feller, *An introduction to probability theory and its applications, vol ii*, John Wiley and Sons, 1971.

[23] J. A. Fill and S. Janson, *The sum of powers of subtree sizes for conditioned Galton–Watson trees*. Unpublished.

[24] J. A. Fill and N. Kapur, *Limiting distributions for additive functionals on Catalan trees*, Theoretical computer science **326** (2004), no. 1-3, 69–102.

[25] B. Haas and G. Miermont, *Scaling limits of Markov branching trees with applications to Galton–Watson and random unordered trees*, The Annals of Probability **40** (2012), no. 6, 2589–2666.

[26] I. A. Ibragimov and Yu. V. Linnik, *Independent and stationary sequences of random variables*, Wolters-Noordhoff Publishing, Groningen, 1971. With a supplementary chapter by I. A. Ibragimov and V. V. Petrov, Translation from the Russian edited by J. F. C. Kingman.

[27] F. Jabot and J. Chave, *Inferring the parameters of the neutral theory of biodiversity using phylogenetic information and implications for tropical forests*, Ecology letters **12** (2009), no. 3, 239–248.

[28] S. Janson, *Stable distributions*, arXiv preprint arXiv:1112.0220 (2011).

[29] ———, *Asymptotic normality of fringe subtrees and additive functionals in conditioned Galton–Watson trees*, Random Structures & Algorithms **48** (2016), no. 1, 57–101.

[30] O. Kallenberg, *Random measures, theory and applications*, Vol. 77, Springer, 2017.

[31] M. Kirxpatrick and M. Slatkin, *Searching for evolutionary patterns in the shape of a phylogenetic tree*, Evolution **47** (1993), no. 4, 1171–1181.

[32] S. Kochen and C. Stone, *A note on the Borel-Cantelli lemma*, Illinois Journal of Mathematics **8** (1964), no. 2, 248–251.

[33] I. Kortchemski, *A simple proof of Duquesne's theorem on contour processes of conditioned Galton–Watson trees*, Séminaire de probabilités xlv, 2013, pp. 537–558.

[34] ———, *Sub-exponential tail bounds for conditioned stable Bienaymé–Galton–Watson trees*, Probability Theory and Related Fields **168** (2017), no. 1-2, 1–40.

[35] J.-F. Le Gall and Y. Le Jan, *Branching processes in Lévy processes: the exploration process*, The Annals of Probability **26** (1998), no. 1, 213–252.

[36] K. Panagiotou and B. Stufler, *Scaling limits of random Pólya trees*, Probability Theory and Related Fields **170** (2018), no. 3-4, 801–820.

[37] J. Pitman, *Combinatorial Stochastic Processes: Ecole d'Eté de Probabilités de Saint-Flour XXXII-2002*, Springer, 2006.

[38] A. F. Poon, *Phylodynamic inference with kernel ABC and its application to HIV epidemiology*, Molecular biology and evolution **32** (2015), no. 9, 2483–2495.

[39] D. Ralaivaosaona, M. Šileikis, and S. Wagner, *A central limit theorem for almost local additive tree functionals*, Algorithmica **82** (2020), no. 3, 642–679. MR4058420

[40] N. Ross and Y. Wen, *Scaling limits for some random trees constructed inhomogeneously*, Electronic Journal of Probability **23** (2018).

[41] J. Scott, P. Maini, A. Anderson, and A. Fletcher, *Inferring tumor proliferative organization from phylogenetic tree measures in a computational model*, Systematic biology **69** (2020), no. 4, 623–637.

[42] K.-T. Shao and R. R. Sokal, *Tree balance*, Systematic Zoology **39** (1990), no. 3, 266–276.

[43] L. Székely, S. Wagner, and H. Wang, *Problems related to graph indices in trees*, Recent trends in combinatorics, 2016, pp. 3–30. MR3526402

[44] G. Voisin, *Dislocation measure of the fragmentation of a general Lévy tree*, ESAIM: Probability and Statistics **15** (2011), 372–389.

[45] S. Wagner, *Central limit theorems for additive tree parameters with small toll functions*, Combin. Probab. Comput. **24** (2015), no. 1, 329–353. MR3318048

Romain Abraham, Institut Denis Poisson, Université d'Orléans, Université de Tours, CNRS, France

E-mail address: romain.abraham@univ-orleans.fr

Jean-François Delmas, CERMICS, Ecole des Ponts, France

E-mail address: delmas@cermics.enpc.fr

Michel Nassif, CERMICS, Ecole des Ponts, France

E-mail address: michel.nassif@enpc.fr